







Db	1264	CTCTTTTGTTGGCTGGCTTCCTGGAGATGCTGCATAAATTTTGTTCCTGCAGCAT	13233
OY	1081	AAAGAAATTCATGACACTTTCGAGAAGGCAGAGCGCCAAAGCTTTGGAGGAAGATATCATC	11440
Db	1324	AAAAAATTCATACACACTTCGAAAGAAAGCGAGAGCGCCAAAGCTTTGGAGGAAGATATCATC	13833
OY	1141	CAAGAAAGTCGTTTCTATTTCCGAGGATATAGCTTTGGGCCACTGCTGACGACGCAAGAGAT	12000
Db	1384	CAAGAAAGTCGTTTCTATTTCCGAGGATATAGCTTTGGGCCACTGCTGACGACGCAAGAGAT	14433
OY	1201	GGAGTCCCAATGGAAGGTTCTGGCATTATATAGTCCCAAGCTCCAGGCCCTTCTTAAGG	12600
Db	1444	GGAGTCCCAATGGAAGGTTCTGGCATTATATAGTCCCAAGCTCCAGGCCCTTCTTAAGG	15033
OY	1261	GAAGGAGAAACACGCGGGAACCTCTAGCTGACATGAGCCAGAGAAATTAACCATCTCC	13220
Db	1504	GAAGGAGAAACACGCGGGAACCTCTAGCTGACATGAGCCAGAGAAATTAACCATCTCC	15633
OY	1321	CGGCGTCTCTTCCCTTTCACCTTTCCTCATCTTCAATATCTTCTACTGGGTTGCTATAAA	13800
Db	1564	CGGCGTCTCTTCCCTTTCACCTTTCCTCATCTTCAATATCTTCTACTGGGTTGCTATAAA	16233
OY	1381	GTCGTATGATGTCAGAAAGATATCCACAGAGCTCTGTGAATAGGCTGGAGCTATAGAGTCT	14400
Db	1624	GTCGTATGATGTCAGAAAGATATCCACAGAGCTCTGTGAATAGGCTGGAGCTATAGAGTCT	16833
OY	1441	GCTCTGAGCCCTCCGCTTCTCCCTGGGAGGCTTTCGCCAGTTAGACTCCATTTAGGG	15000
Db	1684	GCTCTGAGCCCTCCGCTTCTCCCTGGGAGGCTTTCGCCAGTTAGACTCCATTTAGGG	17433
OY	1501	GTTTGGACAGTTCCTTCCCTGATCTCCCACTGAGACTTCAACTACCACTCCCAAGCTAT	15600
Db	1744	GTTTGGACAGTTCCTTCCCTGATCTCCCACTGAGACTTCAACTACCACTCCCAAGCTAT	18033
OY	1561	GTTGGGCTATATTGATAGTGGTCCAAATGCTGCTACTTAATAAGATGCTATCTAACCC	16220
Db	1804	GTTGGGCTATATTGATAGTGGTCCAAATGCTGCTACTTAATAAGATGCTGCTAACCC	18620
OY	1621	TAATAAAAAAAAAA 1632	
Db	1863	TTAATAAAAAAAAA 1874	
RESULT 2			
AAC61678			
ID	AAC61678	standard; DNA; 1600 BP.	
XX	AAC61678;		
AC			
XX			
DT	19-FEB-2001	(first entry)	
XX			
De	cdNA sequence encoding a human ataxia protein.		
XX			
KW	Human; ataxia; gene therapy; ss.		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	1..1254	
FT		/tag="a	
FT		/product="ataxia protein"	
XX			
PN	WO200058461-A1.		
XX			
PD	05-OCT-2000.		
XX			
PF	23-MAR-2000; 2000WO-EP02600.		
XX			
PR	26-MAR-1999; 99EP-0106343.		
XX			
PA	(RAPP/) RAPPOLD-HOERBRAND G.		
XX			
PI	Rappold-Hoerbrand G;		

[illegible]

```
Db 679 GAGAGAGATCTAGGCTGTTGTACCAAGCACTACACAGGCAATTCATCTCATCGAG 738
Oy 781 GTAAGATTTCACCTGAGACGGCAGATGGCTACTATCTGATTCAGATGATACATCCACG 840
Db 739 GTAAGATTTCACCTGAGACGGCAGATGGCTACTATCTGATTCAGATGATACATCCACG 798
Oy 841 CTACTCATGTCATCTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db 799 CTACTCATGTCATCTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 858
Oy 901 CGTGTGGGCTGGGCAACACACGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
Db 859 CGTGTGGGCTGGGCAACACACGATGATGATGATGATGATGATGATGATGATGATGATGAT 918
Oy 961 GCCTCTTTCCTAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
Db 919 GCCTCTTTCCTAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 978
Oy 1021 CTCTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
Db 979 CTCTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1038
Oy 1081 AAGAAATTCATACGACTTCGAGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1140
Db 1039 AAGAAATTCATACGACTTCGAGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1098
Oy 1141 CAAGAAAGTGTCTTATATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
Db 1099 CAAGAAAGTGTCTTATATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1158
Oy 1201 GGAGGCTCAATGGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
Db 1159 GGAGGCTCAATGGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1218
Oy 1261 GAAGGAGAAACACAGCGGAAACTCTACGTGACTGAGCCAGAGAAATTCAGACCATCTCC 1320
Db 1219 GAAGGAGAAACACAGCGGAAACTCTACGTGACTGAGCCAGAGAAATTCAGACCATCTCC 1278
Oy 1321 C-GGGCTGTCTTCCCTTTCATCTTCTGATCTTCAATATCTTCTGCTGCTGCTGCTGCT 1379
Db 1279 CAAGGCGGAATTCCTTTCATCTTCTGATCTTCAATATCTTCTGCTGCTGCTGCTGCT 1338
Oy 1380 AGTGTATGTCAGAGATATCCACAGGCTCTGATGATGAGGAGGAGGAGGAGGAGGAGGAG 1439
Db 1339 AGTGTATGTCAGAGATATCCACAGGCTCTGATGATGAGGAGGAGGAGGAGGAGGAGGAG 1398
Oy 1440 TGTGTGTGCTCTCTGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1499
Db 1399 TGTGTGTGCTCTCTGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1457
Oy 1500 GGTGTTGAGACATGCTTCTGATCTCCACATCAGACATCAATCAACGATCCCAAGCTA 1559
Db 1458 GGTGTTGAGACATGCTTCTGATCTCCACATCAGACATCAATCAACGATCCCAAGCTA 1517
Oy 1560 TGTGTGCTATATGATGATGTCGCAATGTCGCTGATCTTATTAAGATGCTTATCTACC 1619
Db 1518 TGTGTGCTATATGATGATGTCGCAATGTCGCTGATCTTATTAAGATGCTTATCTACC 1577
Oy 1620 CTAAAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1640
Db 1578 CTTAAAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1598

RESULT 3
ID AAQ25722 standard; DNA; 657 BP.
XX
AC AAQ25722;
XX
DT 08-DEC-1992 (first entry)
XX
DE GABA-A receptor beta-subunit.
XX
```

```
KM Expression plasmid; gamma-aminobutyric acid; OmpF; glycine; ss.
XX
OS Synthetic.
XX
PN JP0414683-A.
XX
PD 19-MAY-1992.
XX
PF 05-OCT-1990; 90JP-0267743.
XX
PR 05-OCT-1990; 90JP-0267743.
XX
PA (MITU) MITSUBISHI KASEI CORP.
XX
DR WPI; 1992-214122/26.
XX
PT Prepn. of N-terminal extracellular site protein - by culturing
PT E.coli transformed by a plasmid comprising the tac promoter, ribosome
PT binding site, etc.
XX
PS Claim 1; Page 2; 12pp; Japanese.
XX
CC The sequence given is the gamma-aminobutyric acid (GABA) A receptor
CC beta-subunit. This sequence is used in an expression plasmid operably
CC linked to the tac promoter, a ribosome binding sequence, the E. coli
CC outer membrane protein OmpF signal peptide coding sequence and a
CC sequence coding for the N-terminal extracellular site of either the
CC gamma-aminobutyric acid (GABA) A receptor alpha-subunit or the glycine
CC receptor alpha-subunit. This expression plasmid can be used to
CC transform E. coli to produce an N-terminal extracellular site protein
CC of ion channel direct binding type receptor.
XX
SQ Sequence 657 BP; 171 A; 185 C; 156 G; 145 T; 0 other:
XX
Query Match 22.38; Score 366.2; DB 13; Length 657;
Best Local Similarity 74.28; Pred. No. 1.6e-94;
Matches 511; Conservative 0; Mismatches 133; Indels 45; Gaps 2;
Oy 116 CCCAGCCCATGTCCTGCTGATTCCTAGACAAACTTATGGGCGAATCTGATATG 175
Db 14 CCAGGCTATGTCACTCCCTCGACTTCTGATTAAGCTTATGGAGAGCTTGTGGTATG 73
Oy 176 ATGCCAGATTCGCGCCCAATTTTAAAGGCCACCCGTAAGCTGATGATCTGCAATCTTCA 235
Db 74 ATGCCAGATTCGCGCCCAATTTTAAAGGTCCTCTGTAAGCTGATGATCTGCAATCTTCA 133
Oy 236 TCAACAGTTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 295
Db 134 TCAACAGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 193
Oy 296 GGCAGAGTGAATGACCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 355
Db 194 GGCAGAGTGAATGACCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 253
Oy 356 TCGATCCCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 415
Db 254 TTGACCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 313
Oy 416 GGGCCCACTTCATGAGTGAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 475
Db 314 GGGCCCACTTCATGAGTGAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 373
Oy 476 ATGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 535
Db 374 AGTGTCTTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 433
Oy 536 TCCCATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 595
Db 434 TCCCATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 475
Oy 596 CTCTGCTCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 655
Db 476 -----TTGGTTATACCATGAGCACTCATCTTGTAGTG-- 509
```



```
FT Intron 33976..34114
FT exon /*tag- h
FT exon 34115..34195
FT Intron /*tag- i
FT exon 34196..35759
FT exon /*tag- j
FT exon 35760..35901
FT Intron /*tag- k
FT exon 35902..38781
FT exon /*tag- l
FT exon 38782..38996
FT exon /*tag- m

WO200058461-A1.
XX PD 05-OCT-2000.
XX PF 23-MAR-2000; 2000MO-EP02600.
XX PR 26-MAR-1999; 99EP-0106343.
XX PA (RAPP/) RAPPOLD-HOERBRAND G.
XX PI Rappold-Hoerbrand G;
XX DR WPI; 2000-656166/63.
XX PT Novel nucleic acid sequence encoding human ataxia protein for screening
XX PT compounds useful for treating disorders relating to mutations in ataxia
XX gene
XX PS Claim 6; Page 22-44; 47pp; English.
XX CC The present sequence represents the human ataxia gene. The ataxia
XX CC protein and polynucleotides are useful for diagnosing and treating
XX CC disorders related to ataxia. Ataxia gene sequences are useful in
XX CC gene therapy, and as diagnostic tools or reagents for identifying and
XX CC characterizing genetic defect involved in the disorders and diseases
XX CC related to ataxia.
XX SQ Sequence 39796 BP; 10630 A; 9337 C; 9355 G; 10474 T; 0 other;

Query Match 13.9%; Score 227.8; DB 21; Length 39796;
Best Local Similarity 97.1%; Pred. No. 7.7e-34;
Matches 233; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 263 CCACATGACGACGCGGTGATGCTCTTTCGGGCAACAGTGAATGACCCAGCCTGT 322
DB 33745 CTTCCGAGGACGCTACCGGCTGATGCTCTTTCGGGCAACAGTGAATGACCCAGCCTGT 33804
OY 323 CTTACGAGATATCTGATGACTCTGTGACCTCGATCCCTCATGCTGACTATCT 382
DB 33805 CTTACGAGATATCTGATGACTCTGTGACCTCGATCCCTCATGCTGACTATCT 33864
OY 383 GGAAGCCAGCCTCTTTCGTAATGAGAAAGGGCCAACTCCATGAGCTACCAACG 442
DB 33865 GGAAGCCAGCCTCTTTCGTAATGAGAAAGGGCCAACTCCATGAGCTACCAACG 33924
OY 443 ACAACAGTACTGCGCATCTTCAAGATGGAATGCTGTACAGATCAGGCTGACC 501
DB 33925 ACAACAGTACTGCGCATCTTCAAGATGGAATGCTGTACAGATCAGGCTGACC 33983

RESULT 6
ABI99254
ID ABI99254 standard; cDNA; 2404 BP.
XX AC ABI99254;
XX DT 07-MAR-2002 (first entry)
XX DE Mouse ischaemic condition related cDNA sequence SEQ ID NO:85.
XX
```

```
KW Mouse; Ischaemia; compressive Ischaemia; occlusive Ischaemia;
KM vasospastic Ischaemia; Ischaemic condition; Ischaemic disease; ss.
XX Mus musculus.
XX WO200186188-A2.
XX 22-NOV-2001.
XX PD 18-MAY-2001; 2001MO-JP04192.
XX PR 18-MAY-2000; 2000JP-0145977.
XX (UYN1-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
XX DR WPI; 2002-034733/04.
XX PR P-PSDB; ABB57052.
XX PT Examining the ischemic condition (e.g. occlusive Ischemia) by measuring
XX PT expression levels of particular genes defined in the specification or
XX PT by determining the expression profile of a gene group comprising these
XX PT genes
XX PS Claim 2; Page 265-268; 2690pp; English.
XX CC The present invention describes a method for examining Ischaemic
XX CC conditions, comprising measuring the expression levels of particular
XX CC genes (1) in a test sample or determining the expression profile of a
XX CC gene group in the sample comprising genes selected from (1). The method
XX CC is useful for examining the Ischaemic condition (e.g. compressive
XX CC Ischaemia, occlusive Ischaemia or vasospastic Ischaemia) by measuring
XX CC expression levels of particular genes (ABI99202 to ABI99912, encoding
XX CC the protein sequences in ABB57020 to ABB57374) or by determining the
XX CC expression levels of a gene group comprising these genes. The
XX CC expression levels or expression profiles produced by these genes are
XX CC used as an indicator when screening for Ischaemic condition-Improving
XX CC drugs or therapeutics for Ischaemic diseases. ABI99913 and ABI99914
XX CC represent PCR primers for a mouse Ischaemic condition related sequence,
XX CC which are used in the exemplification of the present invention.
XX SQ Sequence 2404 BP; 705 A; 493 C; 502 G; 704 T; 0 other;

Query Match 13.5%; Score 221; DB 24; Length 2404;
Best Local Similarity 56.5%; Pred. No. 1.5e-52;
Matches 503; Conservative 0; Mismatches 340; Indels 48; Gaps 3;

OY 170 GATATGATGCCAGGATTCGGCCCAATTTAAAGGCCACCGTGAACGTGACCTGCACA 229
DB 279 GTTATGATGCCAGGATTCAGACCAACCTCAAAAGGCAATTCCTGTGATGATGATCA 338
OY 230 TCTTCATCAACATTTTCAGTCCGTCACCAAGCACCAATGACCTACCGGTGATGCT 289
DB 339 TTTTATTAATATTTTGGATTCATTCAAGCAACATGAGCTATAGATTAAACATTT 398
OY 290 TCTTGGGCGCAAGCTGAATGAGACCAAGCCT-GTCTACCGAAGATATCTGA-----TG 343
DB 399 TCTTGAAGACGAATGAGATGAGACCAAGCTACGACTAGTACTTCAAGAGGCTGAC 458
OY 344 ACTCTCTGACCTCGATCCCTCCATGCTGAGCTCTTCTGGAAGCCAGACCTCTTTTG 403
DB 459 ATGCACTGACAGTTGAGACCCACCATCTATATAGTCTTGTGAAACCTGACTATTTCTTG 518
OY 404 CTAATGGAAGAGGCGCACTTCATGAGTGAGCCAGGCAACAAGTATGCTAGCGCATCT 463
DB 519 CAATGAAAGATGCGCAATTTTCATGATGTGACCAAGAAATATCTGTTGTTATATC 578
OY 464 TCAGAGATGGGAATGCTGTACAGCATCAGGCTGACCTCATTTTCTCGCTGATGAG 523
DB 579 TTGGGATGAGAGAGCTGCTGTGAGCATGAGGTTGTTATACATTTGATGCTGCTG 638
OY 524 ACCTCAAGAACTTCCCATGAGCATCCAGACGTGACAGATGACGTTGAGAGCTCATCCA 583
```

Accession	Gene	Protein	Accession	Gene	Protein
D639	ACITTA	ACTGTTTCCCATGGAACACACAGCGTGAATAATGCAACTTGTAGAGCT	639	ACITTA	ACTGTTTCCCATGGAACACACAGCGTGAATAATGCAACTTGTAGAGCT
OY	584	TACTCTGCAGCCCTCTGCCATCTCTGTCACTTTCAGTTGGCTACACCCATGAAAGACCTCG	644	TACTCTGCAGCCCTCTGCCATCTCTGTCACTTTCAGTTGGCTACACCCATGAAAGACCTCG	
D6	693	-----TTGGATATATCAACCGCATGATTTAA	710	-----TTGGATATATCAACCGCATGATTTAA	
OY	644	TGTTTGGAGTGGCTGGAAGATGCTCGCTGTCCCAAGTGGCTGAGGGGCTGACTCGCCCC	703	TGTTTGGAGTGGCTGGAAGATGCTCGCTGTCCCAAGTGGCTGAGGGGCTGACTCGCCCC	
D6	717	GATTCACTGCGCAGTACGAGAGATCCCTGTTCACTTGGAAAAAATTCCTTACCTCAATTTG	777	GATTCACTGCGCAGTACGAGAGATCCCTGTTCACTTGGAAAAAATTCCTTACCTCAATTTG	
OY	704	AGTTTATCTTGGCGGATGAGAAAGATCTAGCGCTGTTGTACCAAGCACTACACACAGGGA	765	AGTTTATCTTGGCGGATGAGAAAGATCTAGCGCTGTTGTACCAAGCACTACACACAGGGA	
D6	777	ATATTTAAAGGAGGATATCGAATATGGCACTGTACAAAATATCTATTAAGGCACCTGGTT	830	ATATTTAAAGGAGGATATCGAATATGGCACTGTACAAAATATCTATTAAGGCACCTGGTT	
OY	764	AATTCACCTCATCGAGGTAAAGTTTCACTGTGAAGGCGAGATGGGCTACTATTCGATTC	822	AATTCACCTCATCGAGGTAAAGTTTCACTGTGAAGGCGAGATGGGCTACTATTCGATTC	
D6	837	ACTACACCTTGTGTGGAGGTCACTTACCTTACGAGACAGAGTGGGTTCTACATGATGG	890	ACTACACCTTGTGTGGAGGTCACTTACCTTACGAGACAGAGTGGGTTCTACATGATGG	
OY	824	AGATGTACATCCCCAGCCTACTCATCAGTCATCCCTGCTGCTGCTTCCTTGGATCAACA	883	AGATGTACATCCCCAGCCTACTCATCAGTCATCCCTGCTGCTGCTTCCTTGGATCAACA	
D6	897	GCCTATATGACCAACCACTTGATGATGTGTTCTCCGCGCTCTTCTTGGATCAACC	950	GCCTATATGACCAACCACTTGATGATGTGTTCTCCGCGCTCTTCTTGGATCAACC	
OY	884	TGATCTGCGCCCTGCGCCCTGTGGGCTGGGCGATCACACCGTGCCTACATGACACCC	944	TGATCTGCGCCCTGCGCCCTGTGGGCTGGGCGATCACACCGTGCCTACATGACACCC	
D6	957	CTGATGCTAGTGGCTGCAAGATCACTTCTGGGATCTTCTCCCTGCTCAGTTGGGCTCAG	1010	CTGATGCTAGTGGCTGCAAGATCACTTCTGGGATCTTCTCCCTGCTCAGTTGGGCTCAG	
OY	944	AGAGCTCTGCGCTCCCGGCGCTCTTTCCTCAAGGTGCTCACTGAGGCAATCGCATCT	1000	AGAGCTCTGCGCTCCCGGCGCTCTTTCCTCAAGGTGCTCACTGAGGCAATCGCATCT	
D6	1017	AGTGCACACACCTTCGCGAGCGGAGCTTCTTAAGTGTCTTATGTGAAGGCGCTGGATGTGT	1070	AGTGCACACACCTTCGCGAGCGGAGCTTCTTAAGTGTCTTATGTGAAGGCGCTGGATGTGT	
OY	1004	GGAATGCTGTGTGCTGCTCTTTTGTGTGCTGCTGCTGCTGATGATGCTG	1054	GGAATGCTGTGTGCTGCTCTTTTGTGTGCTGCTGCTGCTGATGATGCTG	
D6	1077	GGCTCATTTGCTGCTGCTGCTTTCGGGTTGCTCCCTGCTGGAATGACGCTG	1127	GGCTCATTTGCTGCTGCTGCTTTCGGGTTGCTCCCTGCTGGAATGACGCTG	

PT animals comprises the Dermacentor variabilis gamma-aminobutyric acid  
 PT (GABA)-gated chloride channel -  
 XX  
 PS  
 XX Claim 8; Fig 3; 59pp: English.  
 CC The invention relates to gamma-aminobutyric acid (GABA)-gated chloride  
 CC channels and their corresponding nucleic acid molecules. GABA-gated  
 CC chloride channel proteins and DNA's are useful for preventing and  
 CC treating tick infestation, particularly in humans, dogs, cattle, horses,  
 CC deer, or other wild or domesticated animals. The nucleic acids are useful  
 CC as hybridisation probes or Polymerase Chain Reaction primers for  
 CC identifying the presence of Dermacentor variabilis GABA-gated chloride  
 CC channel nucleic acid or DNA's encoding for a GABA receptor. The nucleic  
 CC acids are also useful for the recombinant expression of D. variabilis  
 CC GABA-gated chloride channel proteins. GABA-gated chloride channel  
 CC proteins exert toxic effects on other ticks or related parasites such as  
 CC mites. The present sequence is Dermacentor variabilis clone 8 GABA-gated  
 CC chloride channel DNA.  
 CC  
 XX Sequence 1614 BP; 369 A; 475 C; 422 G; 348 T; 0 other;  
 XQ

Query Match	12.5%;	Score 204.4;	DB 24;	Length 1614;
Best Local Similarity	55.1%;	Pred. No. 7e-48;		
Matches 496; Conservative	0;	Mismatches 356;	Indels 48;	Gaps 31;

Oy	163	ACATCTGGATTTGATATGCGAGGATTCGGCCCAATTTTAAAGGCCCCACCCGTCAACCTGACC	222
Db	154	ACAGTGGGTACGACAGAGGAGGGGTGAGGCCCAATTATGGCGGGCTTCCAGTGGGAATTGGC	213
Oy	223	TGCACATCTTTCATACACATTTTCACTCCGTCAACCAAGACCACATGACTACCCGGGTG	282
Db	214	GTCACATAGCAGATTTATAGCATTAAGTACAGCTCTCTGAAGTACAAATAGCATTTTACTTCT	273
Oy	283	AATGCTCTTATGGCGCAACAGTGGAAATATACCACCGCTGCTTACCGAATATATCTGA -	341
Db	274	GACTTCTATTTCGGCAATCTGGCGGGACAGGACTCTGTTTCCAGAAAAGCCCAAC	333
Oy	342	--TGACTCTGGAACCTCGATCCCTCCATGCTGACTATATCTGGAAGCCAGACCTCTTC	399
Db	334	CTCAGAGACATAGACTGTGGGGCTGMAATGGCCCGAGAGATCTGGGTACCCGACACTTTC	393
Oy	400	TTTGTGTAATGAAAAGGGGCCAACTTCCTATGAGTGACAGGACAGCAAGTAATTTGGGC	459
Db	394	TTCCGCAACGAAAGAGGCGCTACTTTTATGCGGCCACACGCCCAACACTTTCCTCCGC	453
Oy	460	ATCTTCAGAAATGGGAATGCTCTGTACAGATCAGGCTGACCCATTTTGTCTGGCTG	519
Db	454	ATCGGCTCGGAGGAGAGTGTTCGCGAGTATTCACATGAGCGGTACTGCCAGCTGCCCA	513
Oy	520	ATGGACCTCAGAACTTCCCATGAGACATCCAGAGCTGACAGATGACTACCTTGAGAGCTCA	579
Db	514	ATGGATTCACATATCTCCGATGGACAGACAGGCTCAGCTATGAGTTGAAGAAGCT--	571
Oy	580	TGCATACTCTGACAGCCCTCTGCACATCTCTGTCATCTTCAGTTGGCTACACCATGAAGAAG	639
Db	572	-----TTGGTTATACCATGAAGAAGC	591
Oy	640	CTCGTGTTGAGTGGCTGGAAGATGCTCTGCTGTCCAAAGTGGCTGAGAGGGCTGACTCTG	699
Db	592	ATCCGCTATCCGGTGTGCGAGGGTACACCTCCGCTCCCATTCGCCCAAGAGGTTAGAGTTGG	651
Oy	700	CCCCAGTTTATCTTTCGGGATGAGAAGATCTAGGCTGTTGACCAACGACTACAAACACA	759
Db	652	CCGCAAGTTCAAAGTCTCGGTACACTCCAAAAAGC---AAAGAGTTGGCCCTAACGACA	708
Oy	760	GGGAAATTCACCTGCATCGAGGTAAGTTTCACTGGAACGGCAGATGGGCTACTATCTG	819
Db	709	GGAACACTCTCCCGCTGGTATGTGAAATACGCTTGCGCCGCTCCATGGGCTACTACCTG	768
Oy	820	ATTTCAGATGTACATCCCCAGGCTACTCATCGCATCTGTCCTGGGTCTCTTCCTGGATC	879
Db	769	ATCCAGATCTACATCCGGCCGGATGATCGTGTTATTTCTCGGCTCTCTTTTGGCTC	828

OY		880	AACATGATGTGGCCCTGCACC	CGTGTGGCATTACACCGTCACCATGACC	939
Dd		829	CACCGTAGCAGCTTAAGTCCAGCTCCGCTCGCGCTCACACCGTGCTCACCATGACC	888	
OY		940	ACCAGAGCTCTGGCTCCCGGCCCTCTTTGGCTTAAGGTGTCTTCACTGAAGGCAATGCAC	999	
Dd		889	AACTCATGTGTCCACTACACCGACGCGCTGCCCAAATAATCTCAAGTCAAGAATATGCAC	948	
OY		1000	ATCTGGATGGCTGTGTCTGTCTTTGTGTTCGCTGCGCTTGCTGAGATATGTCGCATA	1059	
Dd		949	GTCATACCTGGGCACATGTTTTCTGTAATGATGTTTACCGCGCTCGAGATACCGCGGATA	1008	
<hr/>					
RESULT 8					
ID	AAD22072				
	AAD22072 standard; DNA: 1614 BP.				
XX					
XX	AAD22072;				
DT	12-FEB-2002	(first entry)			
XX					
DE	Dermacentor variabilis clone 5 GABA-gated chloride channel DNA.				
XX					
KW	Gamma-aminobutyric acid; GABA; tick infestation; mite; antiparasitic;				
KM	GABA-gated chloride channel; recombinant expression; domestic animal; ds.				
XX					
OS	Dermacentor variabilis.				
XX					
FH	Key	Location/Qualifiers			
FT	CDS	1..1614			
FT		/*tag=			
FT		/product= "GABA-gated chloride channel protein"			
XX					
PN	MO200174884-A1.				
PD					
XX	11-OCT-2001.				
PE	28-MAR-2001; 2001WO-US09955.				
PR	31-MAR-2000; 2000US-193791P.				
PA	(MERI ) MERCK & CO INC.				
PI	Zheng Y, Cully D, Ludmerer S;				
DR	WPI; 2002-010778/O1.				
DR	P-PSDB; AAE13314.				
PT	New polypeptide useful for preventing or treating tick infestation, in humans, dogs, cattle, horses, deer, or other wild or domesticated animals, comprises the Dermacentor variabilis gamma-aminobutyric acid (GABA)-gated chloride channel -				
PT					
XX					
PS	Claim 8; Fig 3; 59pp; English.				
XX					
CC	The invention relates to gamma-aminobutyric acid (GABA)-gated chloride channels and their corresponding nucleic acid molecules. GABA-gated chloride channel proteins and DNA's are useful for preventing and treating tick infestation, particularly in humans, dogs, cattle, horses, deer, or other wild or domesticated animals. The nucleic acids are useful as hybridisation probes or Polymerase Chain Reaction primers for identifying the presence of Dermacentor variabilis GABA-gated chloride channel nucleic acid or DNA's encoding for a GABA receptor. The nucleic acids are also useful for the recombinant expression of D. variabilis GABA-gated chloride channel proteins. GABA-gated chloride channel proteins exert toxic effects on other ticks or related parasites such as mites. The present sequence is Dermacentor variabilis clone 5 GABA-gated chloride channel DNA.				
CC					
SQ	Sequence 1614 BP; 362 A; 475 C; 429 G; 348 T; 0 other;				
<hr/>					
Query Match	12.5%;	Score 204.4;	DB 24;	Length 1614;	
Best Local Similarity	55.1%;	Pred.No. 7e-48;			

[illegible]





PT DNA encoding alpha-4 and delta subunit(s) of the human GABA  
 PT receptor - also stably co-transfected eukaryotic cells expressing  
 PT receptors contg. these subunit(s), used for screening and designing  
 PT drugs

PS Claim 10; Fig 3; 45pp; English.

XX A cDNA clone (AAT28989) codes for the delta subunit (AAR97299) of the  
 CC human GABA-A receptor. It was isolated from human hippocampus  
 CC cDNA using a probe generated from rat brain cDNA using primers (see  
 CC also AAT28996-97) based on the rat delta sequence. The cDNA can be  
 CC incorporated into a vector and used to direct expression of the  
 CC delta subunit in transfected cells, partic. rodent fibroblasts.  
 CC Co-transfected host cells capable of expressing the delta, alpha-4,  
 CC (see also AAR97298) and beta subunits can be used to screen or design  
 CC drugs that act on GABA-A receptors.

XX Sequence 1555 BP; 289 A; 524 C; 466 G; 276 T; 0 other:

Query Match 11.8%; Score 193; DB 17; Length 1555;

Best Local Similarity 53.3%; Pred. No. 1.3e-44;

Matches 486; Conservative 0; Mismatches 380; Indels 45; Gaps 2;

```

Oy 142 CTGACAACTTATGGGCGGACATCTGGATATGATGCGGAGATTCGGCCCAATTTTAA 201
    || || || || || || || || || || || || || || || || || || || || ||
Db 167 CTCCCAACCTGGAGCGGCTGATAGCCGTTACGCCCACTTCGGCTGGCATCGGA 226
    || || || || || || || || || || || || || || || || || || || || ||
Oy 202 GGGCCACCCGTTGAGAGCTGCAACATCTTCATCAAGTTCAGTTCCTGCTCCGACCAAG 261
    || || || || || || || || || || || || || || || || || || || || ||
Db 227 GGGCCCGGCTGATATGTTGCTTCCCTGGAGGTTGGGACATGACATCATCTCAAG 286
    || || || || || || || || || || || || || || || || || || || || ||
Oy 262 ACCAATGATGATCCGGGTGATGTTCTTGGGGCAACAGTGAATGATCCAGCCCTG 321
    || || || || || || || || || || || || || || || || || || || || ||
Db 287 GCCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 346
    || || || || || || || || || || || || || || || || || || || || ||
Oy 322 TCCTACCGAATATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 381
    || || || || || || || || || || || || || || || || || || || || ||
Db 347 TCTCTAC--AACCAACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 403
    || || || || || || || || || || || || || || || || || || || || ||
Oy 382 TGGAGCCGAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 441
    || || || || || || || || || || || || || || || || || || || || ||
Db 404 TGGCTGCGGAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 463
    || || || || || || || || || || || || || || || || || || || || ||
Oy 442 GACAACAGTATCTGCGGATCTTCAAGATGAGGATGCTGTATACAGATCAGGCTGAC 501
    || || || || || || || || || || || || || || || || || || || || ||
Db 464 GAGACAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 523
    || || || || || || || || || || || || || || || || || || || || ||
Oy 502 CTCATTTTCTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 561
    || || || || || || || || || || || || || || || || || || || || ||
Db 524 TCCACTGTGCTGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 583
    || || || || || || || || || || || || || || || || || || || || ||
Oy 562 ATGAGAGCTGAGAGCTCATCTACTCTGACAGCCCTGCTGCTCTCTCTCTCTCTCT 621
    || || || || || || || || || || || || || || || || || || || || ||
Db 584 CTGGACCTGGAGAGCT-----AC 601
    || || || || || || || || || || || || || || || || || || || || ||
Oy 622 GGCCTACCCATGAAGAGCTCTGTTTGAAGTGGCTGGAAGATGCTCTCTCTCAAGTG 681
    || || || || || || || || || || || || || || || || || || || || ||
Db 602 GGTACTACTCGAGAGATCGTCTACTCTGCTGCTGAGAGAGAGAGAGAGAGAGAG 661
    || || || || || || || || || || || || || || || || || || || || ||
Oy 682 GCTGAGGGCTGACTCTGCGCAGTTTATCTGGCGGAGAGAGAGATCTAGGCTGTCT 741
    || || || || || || || || || || || || || || || || || || || || ||
Db 662 CTGGACAACTGAGTGGCGAGTTCACCATCACCAGTACCGCTTCCACAGAGAGCTG 721
    || || || || || || || || || || || || || || || || || || || || ||
Oy 742 ACCAAGCATCTACACACAGGAAATTCACCTGATGATGATGATGATGATGATGATG 801
    || || || || || || || || || || || || || || || || || || || || ||
Db 722 ATGAATTTCAAGTCGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 781
    || || || || || || || || || || || || || || || || || || || || ||
Oy 802 CAGATGGCTAGTATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 861
    || || || || || || || || || || || || || || || || || || || || ||
Db 782 AACCGGGGCTGATCATCTCAATCTTACATGCTCTGCTGCTGCTGCTGCTGCTGCT 841
    || || || || || || || || || || || || || || || || || || || || ||
Oy 862 TGGGTCTCTCTGATCAACATGATGCTGCGCCCTGCGCGTGTGGGCTGGGCTGATCAC 921
    || || || || || || || || || || || || || || || || || || || || ||

```

```

Db 842 TGGGTCTCTCTGATCAACATGATGCTGCGCCGAGGATGCTCTAGAGCATCAC 901
Oy 922 ACCGTGCTACCATGACACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 981
    || || || || || || || || || || || || || || || || || || || || ||
Db 902 ACGGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 961
    || || || || || || || || || || || || || || || || || || || || ||
Oy 982 TACGTGAGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1041
    || || || || || || || || || || || || || || || || || || || || ||
Db 962 GCCATCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1021
    || || || || || || || || || || || || || || || || || || || || ||
Oy 1042 CTGGAGTATGC 1052
    || || || || || || || || || || || || || || || || || || || || ||
Db 1022 GTGGAGTATGC 1032
    || || || || || || || || || || || || || || || || || || || || ||

RESULT 11
ABK27334
ID ABK27334 standard; DNA; 1359 BP.
AC ABK27334;
XX 09-APR-2002 (first entry)
XX
XX Mutant gamma-aminobutyric acid receptor gamma-2 subunit gene #3.
DE
XX Human; Anticonvulsant; Tranquilliser; Antianemic; Antidepressant;
XX Neuroleptic; Neuroprotective; Neuroleptic; Antimigraine; Anorectic;
XX gamma-aminobutyric acid receptor subunit; GABA; epilepsy; anxiety;
XX manic depression; phobic obsessive symptom; Alzheimer's disease;
XX schizophrenia; migraine; obesity; receptor; gene; ds.
OS Homo sapiens.
XX WO200198486-A1.
XX
XX 27-DEC-2001.
XX
XX 20-JUN-2001; 2001WO-AU00729.
XX
XX 20-JUN-2000; 2000AU-0008260.
XX
XX 13-SEP-2000; 2000AU-000098.
XX
XX 11-MAY-2001; 2001AU-0004953.
XX
XX (BION) BIONOMICS LTD.
XX
XX Wallace RH, Mulley JC, Berkovic SF, Harkin LA, Dibbens LM;
XX WPI; 2002-122280/16.
XX P-PSDB; AAU81289.
XX
XX Mutant gamma-aminobutyric acid receptor subunits and DNA molecule,
XX useful for diagnosing epilepsy, Alzheimer's disease, migraine, obesity,
XX anxiety, manic depression and schizophrenia -
XX
XX Claim 34; Page 73-75; 99pp; English.
XX
XX The invention relates to an isolated mammalian polypeptide (I), which is
XX a mutant of gamma-aminobutyric acid (GABA) receptor subunit. The mutation
XX disrupts the functioning of an assembled GABA receptor, its functional
XX fragment or homologue, and creates a phenotype of epilepsy, anxiety,
XX manic depression, phobic obsessive symptoms, Alzheimer's disease,
XX schizophrenia, migraine and/or obesity. (I), the polynucleotide (II)
XX encoding (I) and antibody (III) to (I) are useful in the diagnosis of
XX epilepsy, anxiety, manic depression, phobic obsessive symptoms,
XX Alzheimer's disease, schizophrenia, migraine and/or obesity. (III) is
XX useful for treating the above conditions. (I)-(III) are useful in
XX screening of candidate pharmaceutical agents, where high-throughput
XX screening techniques are employed. (II) is useful to detect and
XX quantitate gene expression in biological samples. Oligonucleotides or
XX longer fragments derived from (II) are useful as probes in a microarray
XX used to monitor the expression level of large number of genes. (I)-(III)
XX are useful for the study of the function of a GABA receptor, to study the
XX mechanism of the disease as related to GABA receptor, for the creation of.

```





Db 475 TTTCTATG--GAAAGAGGGGATCCTGTACAGGTACACAAAATCTCCACTTGCCACGT 531  
 Oy 706 TTATCTTGGCGGATGAGAGAGATCTAGGCTGTGTCTACCAAGCACTACACAGCGGAAA 765  
 Db 532 TTCA---CGCTGGAAGGTTTCAACGACGACTACGACAGTGGAGCAACAGACTGGCGAG 588  
 Oy 766 TTCACTGCTCATGAGTAAATTTACCTGAGAGGAGGAGTGGCTACTATCTGATTGAG 825  
 Db 589 TACAGCTGCTTGGCGGTGAGCTGGTGTTCAAAGGCGAGTTCAAGCTACTACTGATTCAG 648  
 Oy 826 ATGTACATCCCGACGCTACTCATGTCTCTCTGCTGGTCTCTTGTGATCAACATG 885  
 Db 649 ATTCATCATCCGCTGCTGATCTGTCTCTGCTGGTCTCTGCTGCTGCTGCTGCTGCTG 708  
 Oy 886 GATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 945  
 Db 709 ACCTGATCCCGCGCGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 768  
 Oy 946 AGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1005  
 Db 769 ATATCGGCGATCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 828  
 Oy 1006 ATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1065  
 Db 829 ACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 888  
 Oy 1066 GTTCTCTG 1073  
 Db 889 GCCTCGCG 896

RESULT 14  
 AAD21395  
 ID AAD21395 standard; cDNA; 2138 BP.

AC AAD21395;  
 DT 28-JAN-2002 (first entry)  
 XX  
 DE R. sanguineus glutamate-gated chloride channel 1 cDNA clone, T12.  
 XX  
 KM Brown dog tick; glutamate-gated chloride channel; GluCl1;  
 KM GluCl2; crop protection; insecticide; nematocide; acaricide;  
 KM clone T12; ss.  
 OS Rhipicephalus sanguineus.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 331..1683  
 FT /tag= a  
 FT /product= "R. sanguineus GluCl1 protein, T12"  
 FT /note= "This region is specifically referred in claim 7"  
 XX  
 PR WO200174838-A1.  
 XX  
 PD 11-OCT-2001.  
 XX  
 PF 28-MAR-2001; 2001WO-US09905.  
 XX  
 PR 31-MAR-2000; 2000US-193934P.  
 XX  
 PA (MERI) MERCK & CO INC.  
 XX  
 DR Warmke JW, Yang Y, Cully DF, Hamelin MJ;  
 DR WPI: 2001-662863/76.  
 DR P-PSDB; AAE13037.  
 PT Novel L-glutamate-gated chloride channel proteins from Rhipicephalus  
 PT sanguineus for identifying compounds which modulate the channel  
 PT proteins, which are useful as insecticides, antihelmintics and  
 PT acaricides.

PS Claim 6; Fig 1; 89pp; English.  
 CC The invention relates to Rhipicephalus sanguineus (brown dog tick)  
 CC L-glutamate-gated chloride channel proteins (GluCl1 and GluCl2)  
 CC and nucleic acid molecules encoding such proteins. GluCl1 channel  
 CC proteins are useful for identifying modulators. The compounds  
 CC identified as modulators are useful for insecticidal, mitacidal  
 CC and/or nematocidal treatment for use in animal and human health  
 CC and/or crop protection. The compounds are also useful in screening  
 CC for and selecting compounds active against parasitic invertebrate  
 CC species relevant to animal and human health, including worms,  
 CC fleas, ticks, mites and lice. Heterologous cell lines expressing  
 CC functional GluCl1 and GluCl2 channel functional forms are useful  
 CC for establishing functional or binding assays to identify novel  
 CC GluCl channel modulators. The present sequence is R. sanguineus  
 CC GluCl1 cDNA clone, T12.  
 SQ Sequence 2138 BP; 502 A; 661 C; 561 G; 414 T; 0 other.  
 Query Match 11.4%; Score 187.6; DB 22; Length 2138;  
 Best Local Similarity 55.0%; Pred. No. 5.4e-43;  
 Matches 466; Conservative 0; Mismatches 334; Indels 48; Gaps 3;  
 Oy 226 AACATCTTCATCAACAGTTGAGTCCGTCACCAAGACCAATGAGTACCGGGTAAT 285  
 Db 544 AACATCTTGTAAAGATGATCGGAGAAATGATGACGTACACCATGAGTACACAGTGCAA 603  
 Oy 286 GTCTTTCTGGCGCAACAGTGAATGACCCAGCCCTGCTCTACCCAGAAATATTCGATGAC 345  
 Db 604 ATGACGTTCAAGAGACAGTGGCGGAGAGAGATCCAGATCCAGTACGACTTGGGCGGACG 663  
 Oy 346 TCTCTGACCTCGATCCCTCCATGCTGGAGCTCTATCTGGAAGCCAGACTCTGCTTGTCT 405  
 Db 664 GTTCTCTCTCTGACGCTTACCGAAGCGAGACAGCTTTGGAAGCCGAGACTGTTTCTCC 723  
 Oy 406 AATGAGAAAGGGGCCAATCTCCATGAGGTGACCAAGCAACAAGTACTCGCATCTTC 465  
 Db 724 AACGAGAAAGAGGAGACATCTCCACAAATCATATCATGCCCAAGTCTCTACGATACAT 783  
 Oy 466 AAGAAATGGAATGTCTGTACAGCATGAGGCTGACCCCTCATTTTGTCTGCTGATGAC 525  
 Db 784 CCCAAGGCGAGAGTCTTCTTCAAGCATGAAATATCTGAGCTTTCATGTCGATGAC 843  
 Oy 526 CTCAGAACTTCCCATGAGATCCAGATCCAGATGAGATGAGATGAGATGAGATGAGATG 585  
 Db 844 CTGAATTTTATCTCTTGGATTAACAATCTGCTTATCTGATGATGATGATGATGATGATG 895  
 Oy 586 CTGTCAGCCCTCTGCTGATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 645  
 Db 896 -----ATGGGTATTAACAAGAGAGACTGGTG 921  
 Oy 646 TTTGAGTGTGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 705  
 Db 922 TTTCTATG---GAAAGAGGGGATCTGTACAGGTCAAAAAAATCTCCACTTGGCAGCT 978  
 Oy 706 TTATCTTGGCGGATGAGAGAGATCTAGGCTGTGTTTACCAAGCACTACACAGAGGAAA 765  
 Db 979 TTCA---CGCTGGAAGGTTTCAACGACGACTACGACAGTGGAGCAACAGACTGGCGAG 1035  
 Oy 766 TTCACTGCTCATGAGTAAATTTACCTGAGAGGAGGAGTGGCTACTATCTGATTGAG 825  
 Db 1036 TACAGCTGCTTGGCGGTGAGCTGGTGTTCAAAGCGAGATTGAGCTACTACTGATTCAG 1095  
 Oy 826 ATGTACATCCCGACGCTACTCATGTCTCTGCTGGTCTCTTGTGATCAACATG 885  
 Db 1096 ATTCATCATCCGCTGCTGATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1155  
 Oy 886 GATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 945  
 Db 1156 ACCTGATCCCGCGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1215  
 Oy 946 AGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1005





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: June 30, 2003, 19:53:28 ; Search time 4219 Seconds  
(without alignments)  
11312.776 Million cell updates/sec

Title: US-10-075-846-3  
Perfect score: 1640  
Sequence: 1 atgacaactctgtctcgc.....taaaaaaaaaaaaaaaaaa 1640

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database :

GenBml:.\*  
1: gb\_ba:.\*  
2: gb\_htg:.\*  
3: gb\_in:.\*  
4: gb\_om:.\*  
5: gb\_ov:.\*  
6: gb\_pat:.\*  
7: gb\_ph:.\*  
8: gb\_pl:.\*  
9: gb\_pr:.\*  
10: gb\_ro:.\*  
11: gb\_sts:.\*  
12: gb\_sy:.\*  
13: gb\_un:.\*  
14: gb\_vl:.\*  
15: em\_ba:.\*  
16: em\_fun:.\*  
17: em\_hum:.\*  
18: em\_in:.\*  
19: em\_mu:.\*  
20: em\_om:.\*  
21: em\_or:.\*  
22: em\_ov:.\*  
23: em\_pat:.\*  
24: em\_ph:.\*  
25: em\_pl:.\*  
26: em\_ro:.\*  
27: em\_sts:.\*  
28: em\_un:.\*  
29: em\_vl:.\*  
30: em\_htg\_hum:.\*  
31: em\_htg\_inv:.\*  
32: em\_htg\_other:.\*  
33: em\_htg\_mus:.\*  
34: em\_htg\_pla:.\*  
35: em\_htg\_rod:.\*  
36: em\_htg\_mam:.\*  
37: em\_htg\_vit:.\*  
38: em\_sy:.\*  
39: em\_htgo\_hum:.\*  
40: em\_htgo\_mus:.\*  
41: em\_htgo\_other:.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1521.2	92.8	1874	6 AX392950	AX392950 Sequence
2	1503.8	91.7	1600	6 AX037565	AX037565 Sequence
3	1002.2	61.1	1251	10 AF462147	AF462147 Mus muscu
4	695.8	42.4	2045	5 DRE404970	AJ404970 Danilo rer
5	659.2	40.2	1384	10 RNGRSTR	Y00276 R.norvegic
6	656.2	40.0	1407	10 S73717	S73717 GLRA1-inhib
7	656	40.0	1715	9 HSGLYRA2	X52009 H.sapiens a
8	654.8	39.9	1359	10 RNO310837	AJ310837 Rattus no
9	654.8	39.9	3160	10 RNIGRA2	X61159 R.norvegicu
10	654.8	39.9	3865	10 RNNEOGLY	X57281 Rat NGlyR m
11	654.6	39.9	1350	10 RNO310834	AJ310834 Rattus no
12	654.6	39.9	2125	10 RATGLYRA1	D00833 Rattus norv
13	650.2	39.6	1857	9 HSGLYRA1	X52008 H.sapiens a
14	646.8	39.4	1744	4 AF268375	AF268375 Bos tauru
15	645.4	39.4	1431	10 S73718	S73718 Glra1-inhib
16	643.8	39.3	1374	10 RNO310835	AJ310835 Rattus no
17	643.8	39.3	1679	10 RNIGRA1	X55246 R.norvegicu
18	599	36.5	1222	10 RNO310836	AJ310836 Rattus no
19	585	35.7	2376	5 DRE308516	AJ308516 Danilo rer
20	575.6	35.1	1628	5 DRE5812	AJ005812 Danilo rer
21	572.6	34.9	1418	5 AY094974	AY094974 Morone am
22	552.8	33.7	3069	9 HSU93917	U93917 Human glyci
23	549	33.5	3085	5 AY094975	AY094975 Morone am
24	545	33.2	1443	10 RNO310838	AJ310838 Rattus no
25	543.4	33.1	2413	10 RATGRAS	M55250 Rat inhibit
26	531.2	32.4	1682	5 AF488379	AF488379 Danilo rer
27	510.6	31.1	1392	10 AF362764	AF362764 Mus muscu
28	505.4	30.8	1595	9 DRE308517	AJ308517 Danilo rer
29	498.6	30.4	100269	5 HS1055C14	AL049610 Human DNA
30	366.2	22.3	657	6 E03608	E03608 DNA encodin
31	351.2	21.4	194822	2 AL671887	AL671887 Mus muscu
32	299	18.2	554	6 AX251654	AX251654 Sequence
33	244.8	14.9	2160	5 DRE404971	AJ404971 Danilo rer
34	239	14.6	1491	10 RNO310839	AJ310839 Rattus no
35	235.4	14.4	441	10 AF214575	AF214575 Mus muscu
36	232.6	14.2	1737	9 AF094754	AF094754 Homo sapi
37	232.6	14.2	1737	9 AF094755	AF094755 Homo sapi
38	232.6	14.2	2106	9 HSU33267	U33267 Human glyci
39	232.6	14.2	2649	9 BC032635	BC032635 Homo sapi
40	230.2	14.0	1661	4 AF268376	AF268376 Bos tauru
41	227.8	13.9	39796	6 AX037570	AX037570 Sequence
42	227.8	13.9	39796	9 HSU3563	293848 Human DNA S
43	223	13.6	940	9 AF018157	AF018157 Homo sapi
44	221	13.5	1465	10 MMGRBMRA	X81202 M.musculus
45	221	13.5	2404	6 AX305334	AX305334 Sequence

## ALIGNMENTS

RESULT 1	AX392950	1874 bp	DNA	linear	PAT 23-MAR-2002
LOCUS	AX392950	Sequence 52 from Patent WO0212340.			
DEFINITION	AX392950				
ACCESSION	AX392950				
VERSION	AX392950.1	GI:19700997			
KEYWORDS					
SOURCE	human.				
ORGANISM	Homo sapiens				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.					
REFERENCE	Yue, H., Thornton, M., Ramkumar, J., Tang, Y.T., Azinzal, Y.,				
AUTHORS	Baughn, M.R., Yang, J., Yao, M.G., Lai, P., Wallis, N.K., Ganahl, A.R.,				
	Hafalla, A.J., Nguyen, D.B., Patterson, C., Elliott, V.S.,				



source	1. 1600	/organism="synthetic construct"	
		/db_xref="taxon:32630"	
CDS	1. 1254	/note="cDNA"	
		/note="unnamed protein product"	
		/codon_start=1	
		/transl_table=11	
		/protein_id="CAC16495.1"	
		/db_xref="GI:11226987"	
		/translation="MTLLVPATLISLLLTLPQVILLRVALAKEVSGTSGQPMSP SPLDKLMRGTSYDARIRPNRKGPNVNTCNIFISFSITKTMDYRNVELRNOM NDRLSYREYRSDSLDLPDSMLDSIWKPDLEFANEKGFHEVTITNKLRIKRNW LYSIRLTLISLCLMDLKNFMDIOTCMOLERGYTMKDLVEFMDLDAVOVAEELT LPOEILRDEKDLGCGCTKHNTGKTCIEVFKHLEROMGYLLIOMYLSLIVLISWVS FMTNMDAPARNGIGITVLTMTPTOSSGSRASLPKSYKAKIDIMNAVCLIFPAL EVAITNFSRQHKERIRLRORRORLEEDIIODESRFYRIGELGHLORDGPPMEG SGTSPQPPPLRREGETTRKLYVD"	
BASE COUNT	378 a	440 c	379 g 403 t
ORIGIN			
Query Match	91.7%	Score 1503.8;	DB 6; Length 1600;
Best Local Similarity	96.6%	Pred. No. 0;	
Matches 1385:	Conservative 0;	Mismatches 12;	Indels 44; Gaps 3;
OY	1	ATGACAACCTCTTGTCTGCAACCTCTCTCTTCTCTGAGACCTGCGAGGAG	60
DB	1	ATGACAACCTTCTTCTGCAACCTCTCTCTTCTCTGAGACCTGCGAGGAG	60
OY	61	GTCCTCCTCAGGAGTGGCCCTTGGCAAGAGAGTCAAACTGGAACCAAGGGTCCAG	120
DB	61	GTCCTCCTCAGGAGTGGCCCTTGGCAAGAGAGTCAAACTGGAACCAAGGGTCCAG	120
OY	121	CCCATGTCCCTCTCTGATTTCTAGACAAACTATGGGGGAACTCTGATTTGATGC	180
DB	121	CCCATGTCCCTCTCTGATTTCTAGACAAACTATGGGGGAACTCTGATTTGATGC	180
OY	181	AGGATTGCGGCCCAATTTTAAAGGCCACCGCGTGAAGTGAACATCTTCATCAAC	240
DB	181	AGGATTGCGGCCCAATTTTAAAGGCCACCGCGTGAAGTGAACATCTTCATCAAC	240
OY	241	AGTTTACGCTCGCTACCAAGACCAACAGTACCGGGTGAATGTTCTTGGCGCA	300
DB	241	AGTTTACGCTCGCTACCAAGACCAACAGTACCGGGTGAATGTTCTTGGCGCA	300
OY	301	CAGTGAATGACCCAGGCGTCTACCGAGATATCCGATATCCGATCTCGACCTCGAT	360
DB	301	CAGTGAATGACCCAGGCGTCTACCGAGATATCCGATATCCGATCTCGACCTCGAT	360
OY	361	CCCTCCATGCTGACTCTATCTGGAAGCCAGACCTCTTCTTCTGATGAGAAGGGCC	420
DB	361	CCCTCCATGCTGACTCTATCTGGAAGCCAGACCTCTTCTTCTGATGAGAAGGGCC	420
OY	421	AACTTCATGAGTGAACCAAGCAAAAGTACTCGCATCTTCAAGAAATGGGAATGTG	480
DB	421	AACTTCATGAGTGAACCAAGCAAAAGTACTCGCATCTTCAAGAAATGGGAATGTG	480
OY	481	CTGTACAGATCAGGTGACCTCTATTTGTCTGCTGATGAGACCTCAAGAACTTCCC	540
DB	481	CTGTACAGATCAGGTGACCTCTATTTGTCTGCTGATGAGACCTCAAGAACTTCCC	540
OY	541	ATGGACATCCAGACGTCAGATGATGAGTGAAGTCAATCACTCTGACACCTCTTG	600
DB	541	ATGGACATCCAGACGTCAGATGATGAGTGAAGTCAATCACTCTGACACCTCTTG	600
OY	601	CCATCTCTGCTACCTTCAAGTGGCTACACCATGAAGACCTCGTGTGAGTGGCGAA	660
DB	578	-----TTGGCTTACACCATGAAGACCTCGTGTGAGTGGCGAA	618
OY	661	GATGCTCTGCTGCTGCAAGTGGCTGAGGGGCTGACTCTGCCCGAGTTTACTTCCGGGAT	720
DB	619	GATGCTCTGCTGCTGCAAGTGGCTGAGGGGCTGACTCTGCCCGAGTTTACTTCCGGGAT	678
OY	721	GAGAAGATCTAGGCTGTTGTACCAAGACATACACAGAGAAATTCACCTCATGAG	780
DB	679	GAGAAGATCTAGGCTGTTGTACCAAGACATACACAGAGAAATTCACCTCATGAG	738
OY	781	GTAAGTTTCACCTGGAAGCGGAGAGTGGCTACTATCTGATTCAGATGATCCAGC	840
DB	739	GTAAGTTTCACCTGGAAGCGGAGAGTGGCTACTATCTGATTCAGATGATCCAGC	798
OY	841	CTACTCATGCTCATCTCTGCTCTGGGCTCTCTTGTGATCAACATGATGTGCCCTGCC	900
DB	799	CTACTCATGCTCATCTCTGCTCTGGGCTCTCTTGTGATCAACATGATGTGCCCTGCC	858
OY	901	CGGTGGGCGCTGGGACATACACACCGTGTGCTACCATGACACCCAGAGCTGTGCCG	960
DB	859	CGGTGGGCGCTGGGACATACACACCGTGTGCTACCATGACACCCAGAGCTGTGCCG	918
OY	961	GCCCTCTTGGCTTAAGGTGTCTACGTGAAGGCAATGACATCTGATGCTGTGTCTG	1020
DB	919	GCCCTCTTGGCTTAAGGTGTCTACGTGAAGGCAATGACATCTGATGCTGTGTCTG	978
OY	1021	CTCTTGTGTTGGCTGCTGCTGCTGAGTATGCTGCCATAATTTGTTCTGCTCAGCAT	1080
DB	979	CTCTTGTGTTGGCTGCTGCTGCTGAGTATGCTGCCATAATTTGTTCTGCTCAGCAT	1038
OY	1081	AAAGAAATTCATACGACTTCGAGAAGGACAGGCGCCACGCTTGGAGAAGATATCATC	1140
DB	1039	AAAGAAATTCATACGACTTCGAGAAGGACAGGCGCCACGCTTGGAGAAGATATCATC	1098
OY	1141	CAAGAAAGTGGTTCATATTCCTGTGCTATGAGCTTGGGCCACGCTGACAGGAAGAT	1200
DB	1099	CAAGAAAGTGGTTCATATTCCTGTGCTATGAGCTTGGGCCACGCTGACAGGAAGAT	1158
OY	1201	GAGAGTCCATGGAAGTGTGCTGATTTATAGTCCCACTCAGCCCTCTTCTTAAG	1260
DB	1159	GAGAGTCCATGGAAGTGTGCTGATTTATAGTCCCACTCAGCCCTCTTCTTAAG	1218
OY	1261	GAAGGAAACACACGCGGAAACCTAGTGAAGTGAAGCAAGAAATTCACACATCTGC	1320
DB	1219	GAAGGAAACACACGCGGAAACCTAGTGAAGTGAAGCAAGAAATTCACACATCTGC	1278
OY	1321	C-GGGCTGCTCCCTTTCATCTTCCATCTTCAATATCTTCTACTGAGGTTGCTATAA	1379
DB	1279	CAAGGCGGATTTCCCTTCACTTCCATCTTCAATATCTTCTACTGAGGTTGCTATAA	1338
OY	1380	AGTGCATAGTGTGAGAGATATCCACGAGTGTGAATGGGTGGAGCTATAGAGTCC	1439
DB	1339	AGTGCATAGTGTGAGAGATATCCACGAGTGTGAATGGGTGGAGCTATAGAGTCC	1398
OY	1440	TGCTGCTGGCTCTCTCTCTCTCTGAGGCTTCTCTCCCTCAGTTAGACTCATTAAG	1499
DB	1399	TGCTGCTGGCTCTCTCTCTCTCTGAGGCTTCTCTCCCTCAGTTAGACTCATTAAG	1457
OY	1500	GGTTTGGACAGTTCCTTCTGATCTCCACTCAAGAACTTCACACTACAGTCCCAAAGT	1559
DB	1458	GGTTTGGACAGTTCCTTCTGATCTCCACTCAAGAACTTCACACTACAGTCCCAAAGT	1517
OY	1560	TGTGGGCTTATTTGATGAGTGGCAATGGTGGTGTACTTAAAGATGGCTTATCTACC	1619
DB	1518	TGTGGGCTTATTTGATGAGTGGCAATGGTGGTGTACTTAAAGATGGCTTATCTACC	1577
OY	1620	CTAAAAAAGAAAAAAGAAAAA 1640	
DB	1578	CTAAAAAAGAAAAAAGAAAAA 1598	
RESULT 3			
AF462147			
LOCUS	AF462147	1251 bp	mRNA linear ROD 11-FEB-2002
DEFINITION	Mus musculus glycine receptor alpha 4 subunit (Glr4) mRNA,		
ACCESSION	AF462147		
VERSION	AF462147.1	GI:18448710	
KEYWORDS			

SOURCE Mus musculus.  
ORGANISM Mus musculus.  
REFERENCE 1 (bases 1 to 1251)  
AUTHORS Groemer,T.-W., Becker,C.-M. and Becker,K.  
TITLE Localization of different glycine receptor isoforms in murine spinal cord  
JOURNAL Unpublished  
2 (bases 1 to 1251)  
AUTHORS Groemer,T.-W., Becker,C.-M. and Becker,K.  
TITLE Direct Submission  
JOURNAL Submitted (21-DEC-2001) Biochemistry, University of Erlangen, Fachstrasse 17, Erlangen 91054, Germany  
FEATURES  
source  
1..1251  
/organism="Mus musculus"  
/strain="B6C3/Fe"  
/db.xref="taxon:10090"  
/chromosome="X"  
/tissue\_type="spinal cord"  
1..1251  
/gene="Glr4"  
/note="mmGlr4"  
1..1251  
/gene="Glr4"  
/note="ligand-gated ion channel"  
/codon\_start=1  
/product="glycine receptor alpha 4 subunit"  
/protein\_id="AA16899.1"  
/db.xref="GI:1844871"  
/translation="MSPSDFLDKMGRTSGDARIRNFKPPVNTGFINFSGSV  
ETITMDRVNVEFLKQWNPRLAYREYPDSDLNPSLMSIMKPDLEFNAEKANFH  
EVTIDNKLRIKFNWNLVSLRLILISCPDLNIPMDIOTCTMDESEGYNDMLM  
FEMEDLPAYOVAEGLTLPQILDEKDLGCTHYNTGFTCEVFEHLEROMGYL  
IOMYIPSLIYIISVSWFIMMDAPARVIGITVYLMPTOSGSSASLPKYVYVA  
IDIMMACILEFVPAALLEYAAVNFVSRQNHKFMRLRRORQREDDIIRSRFFFG  
YGLCHCLQARDGEGMESSIYSPPTPLKEGTMRKLYVDRAKRIDTISRAVEPPT  
FLVFNITVWYVKLRSEDIHQAL"  
BASE COUNT 324 a 314 c 281 g 332 t  
ORIGIN  
Query Match 61.1%; Score 1002.2; DB 10; Length 1251;  
Best Local Similarity 87.2%; Pred. No. 8.9e-268;  
Matches 1128; Conservative 0; Mismatches 123; Indels 42; Gaps 1;  
QY 124 ATGTCCCTCTGTGATTTCTAGACAACTATGGGGGGAACATCTGATATGATCCAGG 183  
DB 1 ATGTCCCTCTGTGATTTCTGACAGCTTATGGAGACATCTGATATGATGCTAAG 60  
QY 184 ATTCGGCCCAATTTTAAAGCCCAACCGGTGACGTGACCTGCACATCTTCATCAACAGT 243  
DB 61 ATTAGGCCCAATTTTAAAGGTCCACCTGTGAATGTGACCTGCACATCTTCATCAACAGT 120  
QY 244 TTACGCTCCGTCACCAAGACACAACTAGTACCGGGGTGATGCTTCTTGCGGACAG 303  
DB 121 TTTCGGCTCCGTCACGTGACACCATGACCTACCGGGTAAATGCTTCTTGCGGACAG 180  
QY 304 TGAATGACCCACGCTGTCTTACCGAGAAATATCTGTGACCTCTGACCTGCATTC 363  
DB 181 TGAATGACCCACGCTGTGCTTACCGAGAAATATCTGACCTCTGACCTGCATTC 240  
QY 364 TCCATGCTGACTATCTGAGACCAAGACCTCTTTTGTGTAATGAGAAAGGGCCCAAC 423  
DB 241 TCCATGCTGACTATCTGAGAAACAGACCTCTTTTGTGTAATGAGAAAGGGCCCAAC 300  
QY 424 TTTCATGAGGTGACCAAGACAAAGTACTGGCATCTTCAGAAATGGAATGTGCTG 483  
DB 301 TTTCATGAGGTGACCAAGACAAAGTACTGGCATCTTCAGAAATGGAATGTGCTG 360  
QY 484 TACAGCATGAGGCTGACCTGATTTGCTGCTGCTGATGAGCTCAAGAAATTCCTCCCAT 543  
DB 361 TACAGCATGAGTGAACCTCTATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420

QY 544 GACATCCAGACGTGCAGATGCACCTTGAGAGCTATCATCTGACGCTCTGCA 603  
DB 421 GACATCCAGACGTGCATCATCTGACCTGGAAGCT----- 454  
QY 604 TCTGTGATCTTTAGTTGGCTACACCATGAAAGACCTGCTTTGAGTGGCTGGAAGAT 663  
DB 455 -----TTGGCTACCATGAAATGACCTCATGTTGATGGCTGGAAGAT 498  
QY 664 GCTCTGCTGTCCAAAGTGGCTGAGGGGCTGCTGCGCCGCTTATCTTGGCGGATGAG 723  
DB 499 GCTCTGCTGTCCAAAGTGGCTGAGGGGCTGCTGCGCCGCTTATCTTGGCGGATGAG 558  
QY 724 AAGATCTAGGCTGTGTACCAACACTACCAACAGGAAATTCACCTGCATGAGGTA 783  
DB 559 AAGATCTAGGCTGTGTACCAACACTACCAACAGGAAATTCACCTGCATGAGGTA 618  
QY 784 AAGTTTCACTGAGAGGCGAGATGGGCTACTATCTGATTCAGATGATCCAGGCTA 843  
DB 619 AAGTTTCACTGAGAGGCGAGATGGGCTACTATCTGATTCAGATGATCCAGGCTA 678  
QY 844 CTCATCGTCATCTGTCTGGGCTCTCTCTGATCAACATGATGCTGCGCCGCTG 903  
DB 679 CTCATCGTCATCTGTCTGGGCTCTCTCTGATCAACATGATGCTGCGCCGCTG 738  
QY 904 GTGGGCTGGGCTATCACACCGTCTCACATGACCAACAGACCTGTGCTCCGGGCT 963  
DB 739 GTAGGCTGGGCTATCACACCGTCTCACATGACCAACAGACCTGTGATCCAGGCT 798  
QY 964 TCTTTGGCTAAGGCTGCTCTGATGATGATGATGATGATGATGATGATGATGATGAT 1023  
DB 799 TCTTTGGCTAAGGCTGCTCTGATGATGATGATGATGATGATGATGATGATGATGAT 858  
QY 1024 TTTGTGTTGCTGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1083  
DB 859 TTTGTATTTGCTGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 918  
QY 1084 GAATTCATGACCTTGCAGAGAGGAGAGGCGCCCAACGCTTGGAGAGATATCATCAA 1143  
DB 919 GAATTCATGACCTTGCAGAGAGGAGAGGCGCCCAACGCTTGGAGAGATATCATCAA 978  
QY 1144 GAAAGTGTCTCTTCTTCCGAGGCTATGCTTGGGCGCCACCTGCGAGCAAGAGATGA 1203  
DB 979 GAGAGCGCTTCTTCTTCTTCCGAGGCTATGCTTGGGCGCCACCTGCGAGCAAGAGATGA 1038  
QY 1204 GGTCCAAATGAGAGGCTTGCATTTATAGTCCCAACCTCAGGCTCTTCTTAAGGAA 1263  
DB 1039 GGTCCAAATGAGAGGCTTGCATTTATAGTCCCAACCTCAGGCTCTTCTTAAGGAA 1098  
QY 1264 GGAGAAACCAAGGAGAACTTACGTGACTGAGCCCAAGAGAAATTTGACACCATCTCCGG 1323  
DB 1099 GGAGAAACCAAGGAGAACTTACGTGACTGAGCCCAAGAGAAATTTGACACCATCTCCGG 1158  
QY 1324 GCTGTCTTCCCTTCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1383  
DB 1159 GCTGTCTTCCCTTCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1218  
QY 1384 CTATGTCAGAGATATTCACAGGCTCTGTGA 1416  
DB 1219 CTATGTCAGAGATATTCACAGGCTCTGTGA 1251  
RESULT 4  
DRE404970  
LOCUS DRE404970 2045 bp mRNA linear VRT 21-NOV-2000  
DEFINITION Dnio rerio mRNA for glycine receptor alpha2 subunit (glyr alpha gene)  
ACCESSION AJ0404970  
VERSION AJ0404970.1 GI:11322383  
KEYWORDS glycine receptor alpha2 subunit; glyr gene.  
SOURCE zebrafish.  
ORGANISM Dnio rerio  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;







S73717 1407 bp mRNA linear ROD 17-JUL-2000  
LOCUS GLR1-1 inhibitory glycine receptor alpha 1 subunit [mice, A/HeJ,  
DEFINITION brainstem, spinal cord, mRNA, 1407 nt].  
ACCESSION S73717  
VERSION S73717.1 GI:765206  
KEYWORDS  
SOURCE Mus sp.  
ORGANISM Mus sp.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 1407)  
AUTHORS Ryan, S.G., Buckwalter, M.S., Lynch, J.W., Handford, C.A., Segura, L.,  
Shiang, R., Wasmuth, J.J., Camper, S.A., Schofield, P. and O'Connell, P.  
A missense mutation in the gene encoding the alpha 1 subunit of the  
inhibitory glycine receptor in the spasmodic mouse  
Nat. Genet. 7 (2), 131-135 (1994)  
JOURNAL 95004575  
MEDLINE 7920629  
REMARK Genbank staff at the National Library of Medicine created this  
entry [NCBI g1bbseq 156492] from the original journal article.  
This sequence comes from Fig. 2a.  
FEATURES  
source location/Qualifiers  
1..1407  
/organism="Mus sp."  
/db\_xref="taxon:10095"  
/tissue\_type="brainstem and spinal cord"  
/note="A/HeJ"  
1..1407  
/gene="Glr1"  
58..1407  
/note="alternatively spliced; This sequence comes from  
Fig. 2a; conceptual translation differs from that in  
published reference; GLR1 protein; Glyr alpha 1"  
/codon\_start=1  
/product="inhibitory glycine receptor alpha 1 subunit  
short form"  
/protein\_id="AAB32157.2"  
/db\_xref="GI:9247215"  
/translation="MVSFNTLRFYLMETIVFSLASKAEAPARSPKPMSPDLK  
LMGRTSGDARLRPFRKGPVNVSCNITINSFGSIATETTMQVRIPLRQONDRLA  
TNEYPDSDLDPSMLDSIMKPDLEFANEKGHEHETDNLISRNGVLYSIRI  
TLTLACPMDLKNEPMDVQICIMLESFGYINMDFEIQDQAVQVADGLTLPLPILK  
EKEDRYCTKHYNTGKFTICIEARFLEROMGYLIDYIPSLIYLISMSIWNDA  
APARVGLITTYLTMTTOSSGSRASLPKSYVKAIDIMVAVLLPFLSLEYAVNF  
VSROKELRFRKRHRHHKDEGEGRENFSAVGMPCIAOKDQISVKGANNNTTN  
PPPAKSPPEPRKRFIORAKKIDKISRIQFPMFLIRMFYITKIVREDVANK"  
BASE COUNT 349 a 404 c 332 g 322 t  
ORIGIN  
Query Match 40.0%; Score 656.2; DB 10; Length 1407;  
Best Local Similarity 72.6%; Pred. No. 2,2e-171;  
Matches 942; Conservative 0; Mismatches 298; Indels 57; Gaps 5;

395 TTGACCACTATGTTGGATTCCATCTGGAACCTGACTTGTCTTTGCCAATGGAAG 454  
416 GGGCACTTCATGAGGTGACACGACCAACAACTTACTGGCATCTTCAAGATGGGA 475  
455 GGGCCACTTCACGAATCAGACGAGACACAACTGTAAGATCTCCCGAATGGCA 514  
476 ATGCTCTAGACATCAGAGCTGACCCCTATTTCTCTCTGCTTATGAGACCTCAAGAACT 535  
515 ATGCTCTAGACATCAGAAATCAGACCTGACCTGGCCCTGACCTGAGAAATTT 574  
536 TCCCATGACATCCAGACGTGACAGATGACAGCTTGAAGCTATCATCTCTGACCC 595  
575 TCCCATGATGTATACAGAGCTGTATCATGCAACTGGAAGCT----- 616  
596 CTCTGCAATCTGTCTACCTTTAGTTGGCTACACCATGAACCTCGTGTAGATGAC 655  
617 -----TTGATATACATACAGACCTCATCTTTGAGTGGC 652  
656 TGAAGATCTCTCTGCTCCAGTGGCTGAGGGGCTGACTCTGCCAGTTTATCTTGC 715  
653 AGGACCAAG---GAGCTGTGAGTGGAGGACGAGCTGACCTGCTCAGTTTATTTCTGA 709  
716 GGGATGAGAGATCTAGGCTGTGTACCAACACTACAAACAGAGAAATTCACCTGCA 775  
710 AGGAAGAGAAAGACCTGAGATATCTGACCAACACTACAAACAGGTAAATTCACCTGCA 769  
776 TCGAGTAAAGTTTACCTGAGAGAGAGATGGCTACTATGATGATGATGATGATG 835  
770 TCGAGGCCGATTTCCACCTGAGGAGGAGATGGCTACTACTGATCCAGATGTCATTC 829  
836 CCAGCTACTCTATCTCATCTGTCTGCTGGCTCTCTGATCAACATGATGCTGCC 895  
830 CCAGCTGCTATCTGATCTGCTGCTGATCTCTCTGATCAACATGATGCTGCC 889  
896 CTGCCCCGTGGGCTGGGATCACCACCTGCTCATCATGACCAACCCAGACTCTGGCT 955  
890 CAGCTGTGTGGGGTGGGATCACCACCTGCTCATCATGACCAACCCAGACTCTGGCT 949  
956 CCCGGGCTCTTGGCTAAGGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1015  
950 CCCGAGCTCTCTACCAAGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1009  
1016 GTCTGCTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1075  
1010 GCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1069  
1076 ACATTAAGAAATTCATGACCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1135  
1070 AACCAAGAT 1123  
1136 TCATCCAGAT 1192  
1124 AGGGTGAAGAT 1183  
1193 CAGAT 1249  
1184 CCAAGAT 1243  
1250 CTCTCTTAAGAT 1309  
1244 GGCATTCAGAT 1303  
1310 ACACCATCTCCCGGCT 1369  
1304 ACAAGATATCTGAGATCGGTTTCCCATGAGAGAGAGAGAGAGAGAGAGAGAT 1363  
1370 TTGCTATTAAGTGTATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1406  
1364 TCATCTATTAAGATCGTCCGAGAGAGAGAGAGAGAGAGAGAGAT 1400  
RESULT 7  
HSGLYRA2

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	PUBMED	FEATURES	source
HSLGTYA2	1715 bp mRNA linear	PRI 28-MAY-1993												
H.sapiens alpha-1 strychnine binding subunit of inhibitory glycine receptor mRNA.		X52009.1	GI:31850	glycine receptor; inhibitory glycine receptor; strychnine binding.										
1 (bases 1 to 1715)														
Grnangtjoh.G., Schmieden.V., Schofield,P.R., Seeburg,P.H., Siddique,T., Mohandas,T.K., Becker,C.M. and Betz,H.														
Alpha subunit variants of the human glycine receptor: primary structures, functional expression and chromosomal localization of the corresponding genes														
EMBO J. 9 (3), 771-776 (1990)														
90183975														
2155780														
Location/Qualifiers														
1..1715														
/organism="Homo sapiens"														
/db_xref="taxon:9606"														
/chromosome="X"														
/map="Xp21.2-p22.1"														
/tissue_type="brain"														
/dev_stage="fetal"														
297..1646														
/gene="GLYRA2"														
297..1646														
/gene="GLYRA2"														
/note="strychnine binding alpha-1 subunit"														
/codon_start=1														
/product="inhibitory glycine receptor"														
/protein_id="CA36258.1"														
/db_xref="GI:31851"														
/db_xref="SWISS-PROT:P23415"														
/translation="MISFNTLRLLSGAIVFSSLAASKKEAFAHRSATKPKMSPSFLDK														
LMGPTSDLDARIRNPFGRPPVNVSCNFIENSGSIAETTYMDYRNIFLRQONDPRLA														
LYNEPDDSLDARIRNPFGRPPVNVSCNFIENSGSIAETTYMDYRNIFLRQONDPRLA														
TLTACPMDLNPFMDVOTCIQMLESEGYWNDLIFEOEGAVQVADGLTLPQFIK														
BEKRLRYCTKHNTGKFCIAFPHLEROMGYLYLOWIPLLIVLIVISWIFENWDA														
APAVNGIGITVTLMPTNTOSSGRASLPSKVSVKADIDIMVACILFVSALILEVAVNR														
VSROKELIRRRRRRRHKKDEAGEGRSPNSAYGMCPCADQAKGISYKAKNNSTNTN														
PPPPSSSPPEMKRLFLORAKKIDIKIRISGTFPMAFPLINFTWITIKIVREDVHNO														
BASE COUNT	424 a	480 c	407 g	404 t										
ORIGIN														
Query Match	40.0%	Score 656;	DB 9;	Length 1715;										
Best Local Similarity	71.9%	Pred. No.2.6e-171;												
Matches	952;	Conservative	0;	Mismatches 315;	Indels 57;	Gaps 5;								
OY	89	AGGAAGTCAATCTGGAACCAAGGGGTCACGCCCATCTCCCTTGATTTCCTAGACA	148											
Db	367	AGGAGGCTGAAGCTGCTGCTGCCCAACCAAGCCTATGTCAACCTCGGATTTCCTGATA	426											
OY	149	AACCTATGGGGGGAACATCTGGATATGATGCGCAGGATTCGGGCCCAATTTTAAAGCCAC	208											
Db	427	AGCTAATGGAAGAACTCCGGATATGATGCGCAGGATTCAGGCCCAATTTTAAAGCCACC	486											
OY	209	CCGTGAACGTGACCTGCACAACTCTTCATCAACAGATTTCAGCTCCGTCCACCAAGACACAA	268											
Db	487	CAGTGAACGTGACCTGCACAACTCTTCATCAACAGATTTCAGCTCCGTCCACCAAGACACAA	546											
OY	269	TGCACTACCGGGTGAATGTCTTTTGGGGCAACAGTGGATGAC												

Db	667	CTGACCTGTTTGTGGCCACAGAGAGGGGGCCACTTCCATGAGATCCACAGCAACA	726
Qy	449	AGTTACTGCGCATCTTCAAGAAATGGGAATGTGCTGTACAGCATCAGGCTGCACCTCATTT	508
Db	727	AATTGCTAAAGGATCTCCCGAATGGGAATGTCTCTACAGCATGAGATCAACCCGTGAC	785
Qy	509	TGTCTGCTCTATGACTCAAGAACTTCCCATATGACATCCAGACGTGCACGATGCAGC	568
Db	787	TGGCCCTGGCCCATGGACTTGAAGAAATTTCCCATATGAGATGTCCAGACATGATCATATG	846
Qy	569	TTGAGAGCTCAATCATACTGTGACAGCCCTGTGCATCTGTCACTTCAATTTAGGCTACA	628
Db	847	TGGAAGGT-----TTGGAATTA	864
Qy	629	CCATGAAGACCTGTGTGTTAGTGTGCTGGAAGATGCTCTGTGTGTCCAACTGGCTGAGG	688
Db	865	CGATGAATGACCTCATCTTTGAGTGGCCAGGAACAG---GAGCCGTGCAGGTATACAGATG	921
Qy	689	GGCTGACTCTGCCCGAGTTTATCTTGGCGGATGGAAGAACTATAGGCTGTGTACCAAG	748
Db	922	GACTTACTCTGCCCGAGTTTATCTTGAAGAGAAGAGAACTTGAATATCTGCACCAAGC	981
Qy	749	ACTACACACAGGGAATTCACCTGCATCGAGTMAAGTTTCACCTGGAGCGCAGATG	808
Db	982	ACTACACACAGGTAATTCACCTGCATTTGAGCGCCGGTTCACCTGGAGCGCGAGATGG	1041
Qy	809	GCTACTATCTGATTCAGATGTACATCCCCAGCCTACATCTGATCCTGTCTGAGGACT	868
Db	1042	GTTACTACTATTCAGATGTATATCCCAACCTGTCTATTGTCACTCTCTCATGATCT	1101
Qy	869	CCTTCTGTGATCAACATGGATGCTGCGCCCTGCGCGTGTGGGCTGGGCAATCACACCGTGC	928
Db	1102	CCTTCTGTGATCAACATGGATGCTGCGACTGCTCGTGTGGGCTGAGGATCAACAOTGTGC	1161
Qy	929	TCACCATGACACCCAGACGCTGTGGCTCCGGGCTCTTTGCCCTTAAGGTGTCTTACGTGA	988
Db	1162	TCACCATGACACCCAGACGCTGTGGCTCCAGATCTCTCCCAAGGTGTCTTATGTGA	1221
Qy	989	AGGCAATGCACTGTGATGTGCTGTGTCTCTTGTGTGTGGTCCCTTGTGTGAGT	1048
Db	1222	AAGCATTTGACATTTTGGATGGCAGTTTGGTCCGTCTTGTGTCTCAGCCCTATTAGAT	1281
Qy	1049	ATGCTGCCATTAATTTTGTCTGTGTCAGCATTAAGAATTCATACGACTTGGAAAGAGC	1108
Db	1282	ATGCTGCCATTAATTTGTGTCTGTGGCAACATAAGAGACTCTCCGATTCAGGAGAGAC	1341
Qy	1109	AGAGCGCCCAAGCTTTGGAGAAATATCATCAACGAAGATGCTTTTCAATTTCCGTGACT	1168
Db	1342	GGAGACATCAACAGAGAGATGATGACT-----GGAGAGGCCGCTTTTAACCTTCTGTGCT	1395
Qy	1169	ATGCTGTGGGCC---ACTGCTGCAGCGCAGAGATGAGATGATGATGGAAGTTC---TG	1222
Db	1396	ATGGGATGTGGGCCAGGCTGTCTACAGGCCAAGATGGCATCTCAGTCAAGGGGCCCAACA	1455
Qy	1223	GCAATTTATAGTCCCAACCTCCAGGCCCTCTTGAAGGAAGGGAACCAACCGGGGAAC	1282
Db	1456	ACAGTAACACCAACACCCCTCTCTGACCATCTAACTCCCAAGAGAGATGCGAANAAC	1515
Qy	1283	TCTACGTGTGAGTGAAGCAAGAAATTTGACACCATCTCCCGGGCTGTCTTCCCTTCACTT	1342
Db	1516	TCTTCATCCAGAGGGCCCAAGAAATGACAAATAATATCCCGATTTGGCTTCCCATGTGCT	1575
Qy	1343	TCTCATCTCAATATCTTCTACTGTGGTGTCTATTAAGTCTTATGTCTCAGAGATATCC	1402
Db	1576	TCTCATTTTCAACATGTCTTACTGTGATCATCTACAGATGTCTCGGTAGAGAGACGTGC	1635
Qy	1403	ACCA	1406
Db	1636	ACAA	1639



RESULT 9  
LOCUS RNIGRA2  
DEFINITION R.norvegicus mRNA for inhibitory glycine receptor alpha 2A subunit.  
ACCESSION X61159  
VERSION X61159.1 GI:288344  
KEYWORDS Inhibitory glycine receptor alpha subunit.  
SOURCE Rattus norvegicus.  
ORGANISM Rattus norvegicus.  
Elkayota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE 1 (bases 1 to 3160)  
AUTHORS Kuhse, J., Kuryatov, A., Maulet, Y., Malosio, M.L., Schmieden, V., and Betz, H.  
TITLE Alternative splicing generates two isoforms of the alpha 2 subunit of the inhibitory glycine receptor  
JOURNAL FEBS Lett. 283 (1), 73-77 (1991)  
MEDLINE 91243883  
PUBMED 1645300

FEATURES  
source location/Qualifiers  
1..3160  
/organism="Rattus norvegicus"  
/db\_xref="taxon:10116"  
552..1910  
/codon\_start=-1  
/product="inhibitory glycine receptor alpha 2A subunit"  
/protein\_id="CAA43471.1"  
/db\_xref="GI:288345"  
/db\_xref="SWISS-PROT:P22771"  
/translation="MNRQLYNIIITLAFLEFLGTHNPREACCKDHDSRGKHSQTLSP  
SDFLDKMGRISGIDAHIRNENRGPYNVNTCNIFINISGVLETIMDYRNILRQDM  
NLSRLASEYRPSDLSLDPSMDISIMVPIDLFEANEKANHDVTDNRKLRLIRKRGV  
LYSIRLRLTSLCPMDLKNFPMDOVTQTMQSESGYKMNMDLIFMLSDGPQVAVGLTL  
PQRLKEKEELGYCTKHYNTEKFTCLIEVKEHLEROMGYLIDMAYIPILILVLSWSE  
WIMNDAPARVALGITVLTMTTQSSGSLPKVSVYKVIDIMAVCLIFVPALE  
YAAVNFYSRQHKFEFLRLRRQKRNKEDVRESRFNSGCHQCYVDGTVKAT  
PAMPLEPPPDADAIKKFFVDRAKRIDITISRAFLFLFNFTWITTYKTIIRHEDVH  
KK"

sig\_peptide 552..632  
mat\_peptide 630..1907  
BASE COUNT 902 a 669 c 645 g 944 t  
ORIGIN

Query Match 39.9%; Score 654.8; DB 10; Length 3160;  
Best Local Similarity 70.8%; Pred. No. 6.2e-171;  
Matches 938; Conservative 0; Mismatches 332; Indels 54; Gaps 3;

QY 84 AAAAGAGGAAGTCAATCTGGACCAAGGGGTCCACCCCATGCTCCCTCTGATTTCCT 143  
DB 635 AGACCAATGACTGACGCTGGAAAAACATCCCTGCGACAGCCCTGCTCTTCAGATTTCCT 694  
QY 144 AGACCAATGACTGAGGAGCAACATCTGATATGATGCGAGATTGCGGCCAATTTAAAG 203  
DB 695 GGATTAACCTATATGGAGAGACATCAGATATGATGCAAGAATCAGGCCAATTTTAAAG 754  
QY 204 CCCAGCCGTGAGCTGACGCAACATCTTCATCAACAGTTTCACTGCTCGCTCAACAAGAC 263  
DB 755 TCTCCAGTAAAGCTTACTTGCATATATTTTATCAACAGTTTGGATCGGTGACAGAAAC 814  
QY 264 CACAATGACTACCGGCTGATGCTCTTTCGGGCAACAGTGAATGACCAAGCGCTGTC 323  
DB 815 CACCATGACTACCGAGTGAACAATTTTCTGAGCACACAGTGGAAACGATTACAGCGCTGCG 874  
QY 324 CTAACGAGATATCCGTGATGACTGCTGACCTCGATCCCTCCATGCTGATCTATCTG 383  
DB 875 ATACAGTGAATACCAAGATATTCCTGAGATTGGATTCATCGATGATGATTCGATTGG 934  
QY 384 GAAGCCAGACCTCTCTTCTCTATATGAGAAGGGGCAACTTCATGAGTGAACACGGA 443  
DB 935 GAAACCGATTTGTTCTTTGCGCATATGAGAAGGCAACATTTCCATGATGTCACTGCA 994

QY 444 CAACAACTACTGCGCATCTTCAAGATGGGAATGTGCTGTACAGCATGAGCGTACCT 503  
DB 995 TAACAACTGCTGGGGATTTTCCAAAATGGCAAGTGCTCTACAGATATGACTCACCT 1054  
QY 504 CATTTTCTGCTGCTGATGACCTCAAGAACCTTCCATGAGATCCAGACGTGCAGAT 563  
DB 1055 GACTTATCTGTCGCCATGTGACCGCTGAAGAACTTTCCAAATGATGTCCGACCTGTACAA 1114  
QY 564 GCACCTTGAGAGCGCATCCATFACCTGTCAGACCCCTGTCATCTGTGCACTTTGAGTTG 623  
DB 1115 GCAGCTGAGAGTT-----TTGG 1132  
QY 624 CTACACCATGAAGACCTCGTGTGAGTGGCTGAGAGATGCTCTGCTGCCAAGTGC 683  
DB 1133 GTACACCATGAATGACCTGATATTTGAGTGGTTAAGTATGATGTC---CACTACAGTGC 1189  
QY 684 TGAAGGCTGACTCTGCCCATGTTATCTTCCGGGATGAGAAGATCTAGCTGTGTAC 743  
DB 1190 TGAAGGACTCACCTGCTCAGTTATTTGAAGAAAGAGAAAGAACCTTGCTATTGCGAC 1249  
QY 744 CAAGCATACACACAGGAAATTCACCTGCATGAGGTAAAGTTTCACTGGACGGCA 803  
DB 1250 AAAGCATTAACACACTGSGCAAGTTTACCTGCATTTAGGCTCAAGTTTCACCTGGACGCA 1309  
QY 804 GATGGGCTACTATCTGATTCAGATGTACATCCCAAGCTACATGCTCATCTGTCCTG 863  
DB 1310 GATGGGCTATTTATTTGACAGATGTATATCCCAAGCTCTGATATGATCTATTGTCCTG 1369  
QY 864 GGTCCTCTGATTCATCAATGATGTGCTGCCCTCGCCGCTGGGCTGGGCAATCCAC 923  
DB 1370 GGTCCTCTTTGGATTAACATGATGATGAGTCCGCTCGCAAGGATGCTGCTGGCATCACAC 1429  
QY 924 CTTGCTACCATGACACACCAAGAGCTGCTCGCGGCTCTTTGGCTTAAGTGTCTTA 983  
DB 1430 CGTCTGACATGATGATCCCAAGATTCAGTTCCAGGAGCAATCTGCAAAAGGTCTATA 1489  
QY 984 CGTGAAGCATGACATCTGATGAGTGGCTGTGTCTCTTTGTGTGCTGCTGCTGCT 1043  
DB 1490 TGTGAAGCAATTCATCTGATGAGGAGTGTGCTCTCTGTGTGCTGCTCTTACT 1549  
QY 1044 GAGATATGCGCATTAATTTGTGTTCTGTCAGCATTAAGAAATTCATAGACTGTGAAG 1103  
DB 1550 GGATATGACAGCACTGATCTTTGTCTTCAGGCAACAGAGAGTTCTTGCTGCTCGGAG 1609  
QY 1104 AAGCAGAGCGCGCCACAGCTTGGAGAGATATCATCAAGAAAGTCTTATTTCCG 1163  
DB 1610 ACAGCAAGAGAGGAGATTAAGAGAAAGATGTACTCGTGAAGAACCGTTTAACTTCAG 1669  
QY 1164 TGCGTATGCTTGGCGCACTGCTGCGACGACAGATGAGAGTCAATGAGAGTTCTGG 1223  
DB 1670 CGGTTATGGATGGTCTACTGCTCCAAAGTGAAGATGTTACAGCTGTCAAGGCTTAC--- 1726  
QY 1224 CATTTATAGTCCCAACCTCGAGCCCTCTTCAAGGAGAGGAAACCGAGGAAACT 1283  
DB 1727 -----ACGTGCCAACCCACTTCCGCAACCCCAAGAAAGATGCAAGTCTTCAAGAGAA 1780  
QY 1284 CTACGTGATGAGCAAGAGAAATGACACCATCTCCCGGCTGCTTCCCTTTCACTTT 1343  
DB 1781 GTTTGTGATCGGGCAAAAGAAATCGACACCATATCTCGAGTGGCTCTCCACATGGCTT 1840  
QY 1344 CCTCATCTTCAATATCTTACTAGTGGTGTCTTAAGTCTTATGGTCAAGATATGCA 1403  
DB 1841 CCTCATTTTCAACATCTTTACTGATGATACATCAAGATCATTCGGCATGAAGATGCA 1900  
QY 1404 CCAG 1407  
DB 1901 CAAG 1904

RESULT 10  
LOCUS RNNEOGLY  
DEFINITION Rat NglYr mRNA for neonatal glycine receptor.  
3865 bp mRNA linear ROD 07-AUG-1991

ACCESSION X57281 GI:56743  
VERSION 1  
KEYWORDS glycine receptor; NG1YR gene.  
SOURCE Rattus norvegicus.  
ORGANISM Rattus norvegicus.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
REFERENCE 1 (bases 1 to 3865)  
AUTHORS Hishinuma, F.  
TITLE Direct Submission  
JOURNAL Submitted (28-JUN-1991) F. Hishinuma, Mitsubishi Kasei Institute of Life Sciences, 11 Minamiooya, Machida-shi, Tokyo 194, Japan  
REFERENCE 2 (bases 1 to 3865)  
AUTHORS Akagi, H., Hirai, K. and Hishinuma, F.  
TITLE Cloning of a glycine receptor subtype expressed in rat brain and spinal cord during a specific period of neuronal development  
JOURNAL FEBS Lett. 281 (1-2), 160-166 (1991)  
MEDLINE 91200276  
PUBMED 1707830  
FEATURES  
source Location/Qualifiers  
1..3865  
/organism="Rattus norvegicus"  
/strain="Wistar"  
/db\_xref="taxon:10116"  
/tissue\_type="spinal cord"  
/dev\_stage="10 days old"  
1..3865  
/gene="NG1YR"  
1..3865  
/gene="NG1YR"  
/evidence="experimental"  
1173..2531  
/gene="NG1YR"  
/codon\_start=1  
/product="neonatal type of glycine receptor"  
/protein\_id="CAA0549.1"  
/db\_xref="GI:56744"  
/translating="NMRLVNIILALFAFLGTNHFREARCKDHSRKHPSOTLSP  
SDFDLKMGRTSGDRIAPNFKGPNVAVCNIFINFGSVTEETMTMRYNIFLRD  
NDSRLASEYPPDSDLDPSKMSIDMKPDLFEANERKANPHDVTDKLRISNGYV  
LYSIRLITLITSCPMDFKPNPDVOTCTMQLSEFYTNDLFEWLSGPVQVAGLTL  
PQTLKEKEKELGCTKHNTGKFCICVHLEKROMGYLIOMYIPSLIVILSWVF  
WIMNDAPARVALGELITVTMTTOSGSRASLPKSYVKADIMAMCCLLFVAVALE  
YAAVNFVSRHKEFLRLRRORONKREEDYRESRFNFGMGHCLQVGDGTAVKAT  
PANLPQPPDADAIKKKFDRAKRIIDTISRARPLAFLFNIFWTYTKIRHEDYH  
KK"  
BASE COUNT 1087 a 833 c 889 g 1056 t  
ORIGIN  
Query Match 39.9%; Score 654.8; DB 10; Length 3865;  
Best Local Similarity 70.8%; Pred. No. 6.4e-171;  
Matches 938; Conservative 0; Mismatches 332; Indels 54; Gaps 3;  
OY 84 AAAAGAGAACTCAATCTGGAAACCAAGGGGTCCAGCCCATGTCCCTCTGTATTCCT 143  
DB 1256 AGACCATGACTCCGCTGTGAAACCAATCCCTCGACACCTGTCTCTCAGATTTCCT 1315  
OY 144 AGAACAACCTTATGGGGGGAACATCTGATATGATGCCAGATTTGGGCCAATTTAAAG 203  
DB 1316 GGAATTAACCTAATGGGAAAGACATCAGATGATGACAAAGATCAAGCCAAATTTTAAAG 1375  
OY 204 CCCACCGGTGAACCTGACATCTTCATCAACAGTTTCAGCTCCGTCACCAAGAC 263  
DB 1376 TCCCTCCAGTAAAGCTTACTTCATATTTTATCAACAGTTTGGATCGGTCACAGAAC 1435  
OY 264 CACAATGACTACCGGCTGAATGTCTCTGCGGCAACAGTGAATGACCCACCGCTGTC 323  
DB 1436 CACCATGTACACGAGCAACATTTTCTGAGACACAGTGAACGATTCACGCGTGC 1495  
OY 324 CTACCGAATATATCTGATGACTCTCGAGCTGATCCCTCCATGCTGGAGCTATCTG 383

DB 1496 ATACAGTAGAATCCAGATGATTCCTGTGATTTGGATTCATTCGATTCGATTCGATTTG 1555  
OY 384 GAAGCCAGACCTCTCTCTTGTCTAATGAGAAAGGGGCCAACCTTCATGAGGTGACACGCA 443  
DB 1556 GAACACAGATTTTCTCTTGTGCAATGAGAAAGAGCCAAATTTTCATATGATTCACACTGA 1615  
OY 444 CAACAACTTACTGCGCATCTTCAAGAAATGGAAATGTCCTGTACAGATTCAGGCTGACCT 503  
DB 1616 TAACAAAGTGTGCGGATTTCCAAAATATGGCAAAAGTGTCTGACAGTATTTAGACTCACCTT 1675  
OY 504 CATTTTCTGCTGCTGATGAGACCTCAAGAACTTCCCATGACATTCACAGCTGACAGAT 563  
DB 1676 GACTTTATCTGCTCCATGAGACCTGGAAGAACTTTCCAAATGATGATGTCACACCTGCAAT 1735  
OY 564 GCACCTGAGAGCTATTCATTCATCTGACAGCCCTGCAATCTGTGACATTTGAGTTG 623  
DB 1736 GCACCTGAGAGCTT-----TTGG 1753  
OY 624 CTACACCATGAAAGACCTGTTTGAAGTGAGGCTGGAAGATGCTCTGCTGCAAGTGC 683  
DB 1754 GTACACCATGAAATGACCTGATATTTGAGTGTAAAGTATGATGTC---CAGTACAAATTTGC 1810  
OY 684 TGAAGGCTGACTGTGCCCCAGCTTATCTTGGGGATGAGAAGATATAGGCTTTGTAC 743  
DB 1811 TGAAGGACTCACCTGCTCTGATTTTGAAGAAAGAGAAAGAACTGTGCTATTTCCAC 1870  
OY 744 CAACACCTACAAACAGAGAAATTCACCTGATGAGAGTAAAGTTTCCACCTGGAAGGCA 803  
DB 1871 AAGCATTTACAACTGCGCAAGTTTACCTGATGAGGTCAAGTTTACCTGGAAGGCA 1930  
OY 804 GATGGCTACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 863  
DB 1931 GATGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1990  
OY 864 GCTCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 923  
DB 1991 GCTCTCTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2050  
OY 924 GCTCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 983  
DB 2051 GCTCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2110  
OY 984 GGTGAAGGCAATGACATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1043  
DB 2111 TGTGAAGCAATGACATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2170  
OY 1044 GCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1103  
DB 2171 GCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2230  
OY 1104 AAGCAGAGGCGCCCAACGCTTGGAGAGATATCATCAAGAAAGTGTCTTCTTCTTCCG 1163  
DB 2231 AAGCAGAGAGGCGCAAGATTAAGAGAAAGATGATGATGATGATGATGATGATGATGATG 2290  
OY 1164 TGGCTATGCTTGGGCACTGCTGCAAGCAAGATGAGAGTCCAAATGAAAGCTTCTG 1223  
DB 2291 CGGTTATGAGGATGAGTCACTGCTCAAGTGAAGATGATGATGATGATGATGATGATGATG 2347  
OY 1224 CATTTATAGTCCCAACCTCAGCCCTCTTCTTAAGGAAGAGAAACACGCGGAATCT 1283  
DB 2348 -----ACTGCGCAACCTCTTCCGAAACCCCAAGATGCAATCTTATCAAGAGAA 2401  
OY 1284 CTACGCTGAGTGAACCAAGAAATTTGACACCATCTCCGCGGCTCTTCTCTCTTCACTTT 1343  
DB 2402 GTTTGTGATGAGGCAAAAGAAATGACACCATATCTTCGAGCTGCTTCCACATGCGCTT 2461  
OY 1344 CTTGATCTTCAATATCTTCTTACTGGTGTCTTATTAAGTGTATGATGATGATGATGATG 1403  
DB 2462 CTTGATCTTCAATATCTTCTTACTGGTGTCTTATTAAGTGTATGATGATGATGATGATG 1461  
OY 1404 CCAAG 1407  
DB 2522 CAAAG 2525

RESULT 11  
RNO310834  
LOCUS  
DEFINITION Rattus norvegicus mRNA for glycine receptor alpha 1 precursor,  
1350 bp mRNA linear ROD 15-Aug-2002  
ACCESSION AJ310834.1  
VERSION AJ310834.1 GI:13548654  
KEYWORDS alternative splicing; glycine receptor alpha 1 precursor.  
SOURCE Norway rat.  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE 1  
AUTHORS Beato, M., Groot-Kormelink, P.J., Colquhoun, D. and Sivillotti, L.G.  
TITLE Openings of the rat recombinant alpha 1 homomeric glycine receptor  
as a function of the number of agonist molecules bound  
JOURNAL J. Gen. Physiol. 119 (5), 443-466 (2002)  
MEDLINE 21977760  
REFERENCE 2 (bases 1 to 1350)  
AUTHORS Groot-Kormelink, P.J.  
TITLE Direct Submission  
JOURNAL Submitted (30-MAR-2001) Groot-Kormelink, P.J., Department of  
Pharmacology, The School of Pharmacy, 29/39, Brunswick Square,  
London, WC1N 1AX, UNITED KINGDOM  
COMMENT related splice variants AJ310835 and AJ310836.  
FEATURES  
SOURCE location/Qualifiers  
1..1350  
/organism="Rattus norvegicus"  
/strain="Sprague-Dawley males"  
/db\_xref="taxon:10116"  
/tissue\_type="whole brain"  
/clone\_lib="rat brain marathon-ready cDNA"  
/dev\_stage="8-12 weeks"  
1..1350  
/function="ion channel"  
/note="alternative splicing, primary transcript"  
/codon\_start=1  
/evidence="experimental"  
/product="glycine receptor alpha 1 precursor"  
/protein\_id="CAC35978.1"  
/db\_xref="GI:13548655"  
/translation="MSEFNTLRYLMEITVFPSLAASKREADANSAPKPMSPFLDK  
LMGRTSGDARIRNFKGPVNVSCNIEINSEGSIAETMDYRNIPLRQWNPRLA  
YNEYPDSLDLPMSIDSIKMDLFFANEKGAHFHEITDKRLIRISHNGVLSIRI  
TLTLACPMDLNPFMDVOTCIKMOLESFETNDLIFEQEGAVAVADGLFLPOELK  
EKDLRYCTKHYNTGKFTCIARFHLEROMGVYLIOMYIPLILVILISWISFYNMDA  
APARYGGLITVLTMTTOSSGSRASLRFVSVYKVIDIMAVCLLFVPSALLEYAVNF  
VSRHKELELRFRRRRRHKKDEGEGRNESAYGMGPACLDKDGISVKGANNNTTN  
PAPAPSKSPERMRLFIORAKKIDIKISIFGPMFLIFNMFWIYIKIVREDVHNK"  
BASE COUNT 332 a 395 c 318 g 305 t  
ORIGIN  
Query Match 39.9%; Score 654.6; DB 10; Length 1350;  
Best Local Similarity 72.6%; Pred. No. 6.2e-171;  
Matches 941; Conservative 0; Mismatches 299; Indels 57; Gaps 5;

278 GGCACAGTGGAAAGACACCCCGCTCGCCTACATGATATCTCTGACGACTCTGTGACC 337  
QY  
356 TCGATCCCTCCATGCTGGACTATCTGGAAGCAGACCTCTCTTGGTAATGAGAAG 415  
Db  
338 TTGACCCATCCATGTTGATTCATCTGGAAGCCTGCTTCTTGGCAAGAGAGG 397  
QY  
416 GGGCAACTTCCATGAGTGCACGAGCAACAAGTACTGGGATCTTCCGAATGGA 475  
Db  
398 GGGCCCACTTCCAGATATCCAGCAAGCAAGCAAGAGCTGCTGAGATCTCCCGAGCGCA 457  
QY  
476 ATGTGCTGTACAGCAGCTGACCCCTCATTTTGTCTGCTGATGAGACCTGAGAACT 535  
Db  
458 ACGTCTCTACAGCATATGATATACCTGACTGTGCTGCTGCCCATGAGCTGAGAAAT 517  
QY  
536 TCCCATGACATCCAGACGTCAGATGCAAGCTTGAGAGCTCATCTACTCTGACGCC 595  
Db  
518 TCCGATGATGTACACATATGATATGATGCAAGCTGGAACCT----- 559  
QY  
596 CTCTGCCATCTCTGTACATTTCACTTGTGCTACACATGGAAGACCTGCTTGTGAGTGGC 655  
Db  
560 -----TTGGTATACCATGAAAGAGACTCTATCTTGTGAGTGGC 595  
QY  
656 TCGAAGATGCTCTGCTGTCCAGAGCTGAGGGGCTGACTGTGCCAGTTATCTTTC 715  
Db  
596 ---AAGAGCAAGAGCTGTGAGGTGAGGAGAGGAGCTGACCTGCTCAGATTATCTCTGA 652  
QY  
716 GGGATGAGAGAGATCTAGAGCTGTGTAACCAAGCACTACACACAGGAAATTCACCTGCA 775  
Db  
653 AGGAAGAGAGAGATCTGAGATACATGACACAGCACTACACACAGGTAATTTACCTGCA 712  
QY  
776 TCGAGGTAAGTTCACCTGAGAGGAGAGATGGGCTCTATCTGATTCAGATGATATCC 835  
Db  
713 TTGAGGCCCCGATTCACCTGAGAGAGGAGATGGGCTCTACTACTATCAGATGATATCC 772  
QY  
836 CCAGCTACTATCATGATATCTGCTGTGAGTCTCTCTTGTGATCAACATGAGAGCTGCC 895  
Db  
773 CCAGCTGCTTATGCTGATCTGCTGAGATCTCTCTTGTGATCAACATGAGAGCTGCCAC 832  
QY  
896 CTGCCCCGTGGGCTGGGAGCATACACACGCTGCTCACCATGACCAACCAAGCTCTGGCT 955  
Db  
833 CAGCTGCTGTGGAGCTGGGAGCATACACAGCTGCTCACCATGACCAACAGAGCTCTGGCT 892  
QY  
956 CCGGGGCTCTTGTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1015  
Db  
893 CCGGAGCTCTCTTACCAAGGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 952  
QY  
1016 GTCTGCTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1075  
Db  
953 GCTGCTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1012  
QY  
1076 AGCATAGAATTCATACGATCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1135  
Db  
1013 AACCAAGAGAACTC-----CTTGATTTAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1066  
QY  
1136 TCATCAGAGAAAGTCTTATTTCTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1192  
Db  
1067 AGGTGGAGAAAGCCGCTTCAACTTCTCCGCTATGAGATGGGCCAGGCTGCTGCAAG 1126  
QY  
1193 CAAGATGAGAGTCCATGAGAGGCTTGGCACTTATATAGTCCCAACCTCCAGCCCT- 1251  
Db  
1127 CCAAGATGAGATCTCTGCTCAAGGCTGCAACACACACACACACACACACACACACACAC 1186  
QY  
1252 --CTTCTAAGGAGAGAGAAACACAGCGGAAACCTCTAGTGGATGAGAGGAGGAGGAG 1309  
Db  
1187 CACCGTCCAAATGCCCGGAG 1246  
QY  
1310 ACACCATCTCCCGGCTGTCTTCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1369  
Db  
1247 ACAAAATATCTGCACTGCTTCCCATGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1306  
QY  
1370 TTGCTATTAAGTGTCTATGCTCAGAGATATCCACA 1406  
Db  
1307 TCATCTACAGATGCTCGGAG 1343



RESULT 12  
RATGLYRA1 2125 bp mRNA linear ROD 16-MAY-1998  
LOCUS Rattus norvegicus Glyralphal mRNA for glycine receptor alpha 1,  
DEFINITION complete cds.  
ACCESSION D00833  
VERSION D00833.1 GI:220750  
KEYWORDS glycine receptor alpha 1; GlyRalphal.  
SOURCE Rattus norvegicus (strain:Mistar) older than 25 days spinal cord  
cDNA to mRNA.  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE 1 (sites)  
AUTHORS Grenningloh, G., Rianitz, A., Schmitt, B., Methfessel, C., Zensen, M.,  
TITLES Bayreuther, K., Gundelfinger, E.D. and Betz, H.  
JOURNAL The strychnine-binding subunit of the glycine receptor shows  
MEDLINE homology with nicotinic acetylcholine receptors  
REFERENCE Nature 328 (6127), 215-220 (1987)  
AUTHORS 2 (bases 1 to 2125)  
TITLES Aaga, H., Hirai, K. and Hishinuma, F.  
JOURNAL Functional properties of strychnine-sensitive glycine receptors  
MEDLINE expressed in Xenopus oocytes injected with a single mRNA  
REFERENCE Neurosci. Res. 11 (1), 28-40 (1991)  
AUTHORS 3 (bases 1 to 2125)  
TITLES Hishinuma, F.  
JOURNAL Direct Submission  
MEDLINE Submitted (28-JUN-1991) Fumio Hishinuma, Mitsubishi Kasel Institute  
REFERENCE of Life Sciences, Department of Molecular Biology; 11 Minamiooya,  
AUTHORS Mechida, Tokyo 194, Japan (Tel:0427-24-6233, Fax:0427-25-1252)  
JOURNAL Location/Qualifiers  
FEATURES  
source  
1..2125  
/organism="Rattus norvegicus"  
/strain="Mistar"  
/db\_xref="taxon:10116"  
/tissue\_type="spinal cord"  
/dev\_stage="older than 25 days"  
/note="314 bp upstream of HindIII site"  
1..2125  
/gene="GlyRalphal"  
185..1534  
/gene="GlyRalphal"  
/codon\_start=1  
/product="glycine receptor alpha 1 precursor"  
/protein\_id="BAA0707.1"  
/db\_xref="GI:220751"  
/translation="MYSFNLRFLYMETIVEFSLAASKEDAAARSAKPMSPDFLAK  
LMGRTSGDARIRNPFKSPVNSCNIFNSPFIATDTQRYNITLROOMDPRIA  
YNEYRDSLDLSDMSLDSIKPDLFANEGARHETTTNKLIRISNGNVLSITAI  
TLVLACPMDLKRFPMVDQTCIMLESFVTMNDLIFEMODGAVQADGLTLPQFLIK  
EERDLRYCTHYNTGKFTCLFARHLEKRGYLIOMYIPSLVILVLSFWLMDA  
APRNVGLITFTVLTMTOSGSRASLPKVSVAKIDIMAVQCLIFVSALLEVAANF  
VSRQKELLFRFRKRHRHKKDEGEGEFNSAAGMPCAOAKGSIKVGKNNNTTN  
PAPAPSKSPSEMRKLFIQRAKKIDKISRIGPMAFLFNNEWIIYKIVREDVHNK"  
185..268  
/gene="GlyRalphal"  
/gene="GlyRalphal"  
269..1531  
/product="glycine receptor alpha 1"  
BASE COUNT 536 a 593 c 527 g 469 t  
ORIGIN  
Query Match 39.9%; Score 654.6; DB 10; Length 2125;  
Best Local Similarity 72.6%; Pred. No. 6.6e-171;  
Matches 941; Conservative 0; Mismatches 299; Indels 57; Gaps 5;  
OY 116 CCGAGCCATGTCCTCGATTTCTAGACAACTTATGGGCGCAACATCTGGATATG 175  
II IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII

Db 282 CCAAGCCTATGACACCTCGGACTTCCTGGATTAAGCTTATGGGAAGACTTCTGGATG 341  
OY 176 ATGCCAGATTCGGCCCAATTTAAAGCCACCACCGAGACGATCTGCAACATCTCA 235  
Db 342 ATGCCAGATTCAGACCCCACTTAAAGTCTCTCTGGAACGATGATGCAACATCTCA 401  
OY 236 TCAACAGTTTCAGCTCCGTCACCAAGACACAAATGAGTACCGGGTAAATGCTTCTTGC 295  
Db 402 TCAACAGCTTTGGTTCATATGCGCCGAGACACATGAGTACAGGGTCAACATCTTCTGA 461  
OY 296 GGCACAGTGAATGACCCACGCTGCTACCGAGAATATCTGATGATCTCTGGACC 355  
Db 462 GGCAGCAGTGAAGACGACCCCGTCCGCTACATGATATCTGAGACGATCTCTGGACC 521  
OY 356 TGCATCCCTCATGCTGAGCTCATATGGAAGCAGACGCTCTTGTGTAATGAGAAG 415  
Db 522 TTGACCTCATGATTTGGATTCCATCTGGAAGGCTGATCTTGTTCGAATGAGAAG 581  
OY 416 GGGCAACATTCATGAGTGCACGACGACAAACAGTACTGCCATCTTCAAGAAATGGA 475  
Db 582 GGGCCACTTCCACGAGATCACCACGACAAACAAAGCTGTGAGATCTCCGGAGCGCA 641  
OY 476 ATGTGCTGTACAGCATAGGCTGACCTCATTTTGTCTGCTGATGAGACCTCAAGACT 535  
Db 642 ACCTGCTGTACAGCATAGGATCACCCTGACTGCGCTCCCATGAGACCTCAAGAAAT 701  
OY 536 TCCCATGAGACATCCACAGCGACGATGAGCTGAGAGCTCATCATCTGACGACC 595  
Db 702 TCCCATGAGACATCCACAGCATATGATGCACTGGAAGCT----- 743  
OY 596 CTCTGCACTCTCTCTCACTTTCAGTTGGCTACACATGAAAGACCTGTTGAGTGGC 655  
Db 744 -----TTGGTTATACCATGAAGACGACCTCATCTTTGAGTGGC 779  
OY 656 TGGAAAGTCTCTCTGCTGTCCAAAGTGGCTGAGGGGCTGACTCTGCCCATTTATCTTGC 715  
Db 780 ---AAGAGCAAGAGAGCTGTGAGTGCAGATGAGTGCACCTGCTCCAGTTTATCTCTGA 836  
OY 716 GGGATGAGAGAGATCTAGGCTGTGTACCAAGCACTCAACACAGGAAATTCACCTGCA 775  
Db 837 AGGAAGAGAAAGATCTAGATACGTACACACCAAGCACTCAACACAGGTAATTCACCTGCA 896  
OY 776 TCGAGTAAAGTTTCACCTGGAACGAGATGGGCTACTATCTGATTCAGATGATATCC 835  
Db 897 TTGAGGCCCATTCACCTGGAACGAGATGGGCTACTACGTGATCCAGATGATATCC 956  
OY 836 CCAGCCTACATCATGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 895  
Db 957 CCAGCCTGCTTATCTGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1016  
OY 896 CTGCGCCTGAGGCTGAGGCTGATCCACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 955  
Db 1017 CAGCTGCTGAGGCTGAGGCTGATCCACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1076  
OY 956 CCGGCGCTCTTGGCTTAAGGTGCTACGTGGAAGCAATGACATCTGATGAGCTGTGT 1015  
Db 1077 CCCGAGCTCTCTACCAAGGTGCTACGTGGAAGCAATGATGATGATGATGATGCTGCTT 1136  
OY 1016 GTCTGCTCTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1075  
Db 1137 GCTGCTCTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1196  
OY 1076 AGCATTAAGAAATTCATGAGTTCGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGATG 1135  
Db 1197 AACACAGAGAACTC-----CTTGATTTTATGAGAGAGAGAGAGAGAGAGATG 1250  
OY 1136 TCATCAAGAAAGTCTGTTCTATTTCCGTGCTATGCTTGGGCT---ACTGCTGAGG 1250  
Db 1251 AGGTGAG 1250  
OY 1193 CAGAGATGAGAGTCAATGAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1250  
Db 1311 CCAAGATGAGATCTCTGCTCAAGAGGTGCTCAACACACACACACACACACACACACACAC 1311

[illegible][illegible]

QY	DB	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	TITLE	JOURNAL	AUTHORS	FEATURES	gene	CDS
0y	1263	AGGAAACCCGCGGAAACCTCTACGGGATGAGCCAAAGAAATGACACCATCTCCG	1322	AF268375	1744 bp	mRNA	linear	MAM 22-MAY-2001							
0y	1396	TGGAGATGCTATCAAGAAGATTGTGGACGGGCAAAAGATGACAGATATCTCG	1655	AF268375	1744 bp	mRNA	linear	MAM 22-MAY-2001							
0y	1323	GCGTCTCTCCCTTCACTTCTCTCACTTCCTCAATATCTTCTACGCGGTTGCTATTAAGT	1382	AF268375	1744 bp	mRNA	linear	MAM 22-MAY-2001							
0y	1656	AGCTGCGCTCCCAATTTGGCGCTTCCTCCTATTTTCAACATCTTTTACAGATCAATCAAGAT	1715	AF268375	1744 bp	mRNA	linear	MAM 22-MAY-2001							
0y	1383	GCTATGCTCAAGATATCCACCG	1407	AF268375	1744 bp	mRNA	linear	MAM 22-MAY-2001							
0y	1716	CATTCGGCATGAAATGTCACCAAG	1740	AF268375	1744 bp	mRNA	linear	MAM 22-MAY-2001							
0y	89	AGGAATCAATCTGGAACCAAGGATCCAGCCCATGTCCCCCTCTGATTTCTTAAGCA	148	AGGAATCAATCTGGAACCAAGGATCCAGCCCATGTCCCCCTCTGATTTCTTAAGCA	148	mRNA	linear	MAM 22-MAY-2001							
0y	299	AGGAGCGTGAAGCTGCTCGCTCTCTCAAGCCCATGTCACGCTCGATTTCTTGATA	358	AGGAGCGTGAAGCTGCTCGCTCTCTCAAGCCCATGTCACGCTCGATTTCTTGATA	358	mRNA	linear	MAM 22-MAY-2001							

[illegible]

QY	1208	CAATGGAAGCTC---TGCGATTATATAGTCCCAACCTCCAGCCGCCCTCTTAAGGGAG	1266
Db	1394	CTGTCAAGGGCCCAACAACAGCAACAGCAACCCCTCTCCGGACCATTCACAGTCC	14533
QY	1265	GAGAAACCAACCGGAAACTCTACGTGTGACTGACGCCAAGAGATTGACACCATCTCCGGG	13242
Db	1454	CGAGAGAGATGCGAAAACTCTTCATCCAGGGGGGCAACAGATGACAAAGATCTCCGCCA	15133
QY	1335	CTGTCTTCCCTTTTCACTTTCTCTCATCTTCAATATCTTCTACTGGGTGTCTATAAAGTGC	13844
Db	1514	TCCGGTTCTCCCATGGCCCTTCTCTCATCTTCAACATGTTCTACTGATCATCTAACAAATCG	15737
QY	1385	TATGTCTAGAAATATCCACCA	1406
Db	1574	TCCGACGGAGGAGCTCCACAA	1595
RESULT	15		
LOCUS	S73718		
DEFINITION	S73718	1431 bp	mRNA
ACCESSION	S73718		
VERSION	S73718		
KEYWORDS	S73718		
SOURCE	S73718	GI:765208	
ORGANISM	Mus sp.		
REFERENCE	Mus sp.		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 1431)		
TITLE	Ryan, S.G., Buckwalter, M.S., Lynch, J.W., Handford, C.A., Segura, L., Shiang, R., Wasmuth, J.J., Camper, S.A., Schofield, P., and O'Connell, P. A missense mutation in the gene encoding the alpha 1 subunit of the inhibitory glycine receptor in the spasmodic mouse		
JOURNAL	Nat. Genet. 7 (2), 131-135 (1994)		
MEDLINE	95004575		
PUBMED	7920629		
REMARK	GenBank staff at the National Library of Medicine created this entry [NCBI gidsq 156497] from the original journal article. This sequence comes from Fig. 2a and 2b.		
FEATURES			
source	1. 1431		
gene	/organism="Mus sp."		
CDS	/db_xref="taxon:10095"		
	/tissue_type="brainstem and spinal cord"		
	/note="A/HeJ"		
	1. 1431		
	/gene="Glr1"		
	58. 1431		
	/gene="Glr1"		
	/note="alternatively spliced: This sequence comes from Fig. 2a and 2b; conceptual translation differs from that in published reference; GLR1 protein; Glyr alpha 1"		
	/codon_start=1		
	/product="inhibitory glycine receptor alpha 1 subunit long form"		
	/protein_id="AAB32158.2"		
	/db_xref="GI:9247216"		
	/translation="MYSEFNTLRFYLMETVIFESLSAASKAEARSAPKSPSPFLDK LMGTSQGYDARIIPNFKGPPYVNCNFINFSFGSIATYTDYRNITLROONDPRILA YNEPDSLDLSPSDLSIMKPDLPFRNENGAHHEITTDKLRISRNGVLTYSIRI TLTLACPMDLNPFMDVQTCIMQLESFGYTMNDLIFPMODGAOVADGTLTPOFILAK EKDRLGCTHYNTGKFTCIEARHLEROMGYLIDMYIPSLILVSWISFWINDMA APAVAGCTITVLTMTQSSGRASLRKYSVKAIDIMMVCILFVSALEVAANFV VSRHKELRRRKRHRHKSPLMLNFDODDGGEGEERFMSVAGMGPALOKKDGISVNG ANNNTPRPAPSPKSPBEAKLEFIQAKKIDIKTSIRIGFPMALIFIMFYIITKIVR REDVHNK"		
BASE COUNT	355 a	411 c	336 g
ORIGIN			329 t
Query Match	39.4%	Score 645.4	DB 10; Length 1431;
Best Local Similarity	71.9%	Pred. No. 2.3e-168;	

	Matches	945;	Conservative	0;	Mismatches	301;	Indels	69;	Gaps	5;
QY	116	CCACAGCCATGTC	CCCCCTGTGATTTTCT	AGACAACTTATGGGG	GAACATCTGATATG	175				
Db	155	CCAAAGCTATGTC	ACCTCGGACCTTC	GTGGATTAAGCTCATGGG	AGACCTCTGTGGATATG	214				
QY	176	ATGCCAGATTCGG	CCCAATTTTAAAGGCC	CAACCCGTACAGTGA	CTGCAACATCTTCA	235				
Db	215	ACGCCAGATCAGG	CCCACTTTAAAGGTC	CTCCTGTAGAGTAA	GGTTGCAACATCTTCA	274				
QY	236	TCAACAGTTTACG	CTCGGTCAACCAAG	ACACACATGGATAC	CGGGTGAATGCTTTCTGC	295				
Db	275	TCACAGTTTGCTT	CAATCGCTGAGAAC	ACCATGGATATAGGGT	TCAACATCTTCTCGA	334				
QY	296	GGCAACAGTGAAT	GAACCCACGCGCTCT	ACCGAATATCTGAT	ACTCTGTGACG	355				
Db	335	GGCAGACATGGA	AGACCCCGTCTGGCT	ACAATGATACCTGAT	ACTCTGTGACG	394				
QY	356	TCGATCCCTCAT	GTGTGACTCTATCT	GGAAGCAGACCTCT	CTTTGCTAATGAGAAG	415				
Db	395	TTGACCCCATCT	ATGTTGATTCATCT	GGAAGCCGACTTGTCT	TTTCCATGAGAAG	454				
QY	416	GGGCAACTTCGAT	AGAGTGAACGACCA	CAAGATTACTCGCAT	CTCTCAAGAATGGA	475				
Db	455	GGGCCCCATTC	ACCAAAATACAC	CGACAACAACTCTA	AGAAATCTCCGGATGGA	514				
QY	476	ATGTGCTGTACAG	TCAGGCTGACCCCT	CAATTTTGTCTCGCT	GATGAGACTTCAAGACT	535				
Db	515	ATGTGCTGTACAG	TCAGATCAAGAAC	CCCTGACGCTGGCT	GTGCCCATCTGAAGATT	574				
QY	536	TCCCATGTGACAT	CCACAGCGTGCAC	AGATGCAAGCTTGAG	AGCTCATCTGTCGACCC	595				
Db	575	TCCCGATGAGAT	GTACAGACGTGTAT	CAATGCAACTCGAAAGCT	-----	616				
QY	596	CTGTGCCATCT	CTGTGACTTCA	GTGGCTACACATAA	GAAGACTGTGTTAGTGGC	655				
Db	617	-----	-----	TTGATTAACCAT	TGAAGACTCTATCTTTAGTGGC	652				
QY	656	TGGAAGATGCT	CTCTGCTGTCC	AGTGGGTGAAGGG	GTACTCTGCCCATTTATCTTGC	715				
Db	653	AGGAGCAAG---	GAGCTGTGCAGGT	GGGACAGCGGACT	GTGACCTGTGATTTATTTCTGA	709				
QY	716	GGGATGAGAGAT	CTCTAGCTGTGT	ACCANAGCACTAC	AACAAGGAATTCACCTGCA	775				
Db	710	AGGAAGAAAGAC	CTGATGATCTGAC	CAACAGCACTAC	AACAAGGTAAATTCACCTGCA	769				
QY	776	TCGAGGTAAATTC	ACCTGGAAGGGCA	AGGGGCTACTATCT	GATTCAGATGTACATCC	835				
Db	770	TCGAGGCCCATTC	ACCTGGAAGGGCA	AGGGGCTACTATCT	GATTCAGATGTACATCC	829				
QY	836	CCAGCTCTACTAT	CGTCACTCTGTCT	GTGGTCTCTCTT	GTGATGAACAATGGATCTGCC	895				
Db	830	CCAGCTCTCATGT	CTCATCTGTCTGT	GTGGATCTCTT	GTGATGAACAATGGATCTGTAC	889				
QY	886	CTGCCCGTGTGG	CGCTGGGCATCAC	ACACCGTGTAC	ACCATGACACCCAGAGCTGTGGCT	955				
Db	890	CAGCTCGTGTGG	CGCTGGGCATCAC	ACAGTGTAC	ACCATGACACAGAGCTGTGGCT	949				
QY	956	CCCCGGGCTCTT	GGCTTAAGGTGTCT	TAACGGAAGCAAT	GCAGATCTGGATGGCTGTGT	1015				
Db	950	CCCCGAGCTCT	CCCTACCAAGGTGTCT	TAACGGAAGCAAT	GCAGATTTGGATGGCTGTGT	1009				
QY	1016	GTCGTGCTTTGT	GTGTGCTGCTCC	TTGGTGGAGAT	ATGCTGCCATTAATTTTCTGCTC	1075				
Db	1010	GCTGTGCTTTGT	GTGTGCTGCTCC	CTGTGGAGAT	ACGCCCGTCAACTTTTGTCTGGC	1068				
QY	1076	AGCATAAAGAT	TTATACGACTTGAAG	-----	AAAGCAGAGCGGCC	1117				
Db	1070	AAACACAAGAC	TCGCTTCGATTTAG	GAGGAAGAGGCG	CAATCACAGAAGGCCCATGTCTAA	1129				
QY	1118	AACGCTTGGAG	AAGATATCATCA	AGAAGATGCTTTCT	ATTTCCGTGGCTATGCGCTTGG	1177				
Db	1130	ATCTGTTTACAG	ATGATGAGGGGTG	AGAGGCCCTTCA	ACTTCTGTGCTTATGGAATGG	1189				



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: June 30, 2003, 22:23:50 ; Search time 3552.13 Seconds

(without alignments)  
3531.204 Million cell updates/sec

Title: US-10-075-846-4  
Perfect score: 2252  
Sequence: 1 MTTVPATLFLSLMLTLPLGQ.....PQPPAPLREGETTKLYVD 431

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-O=/cgn2.1/USPRO/US10075846/runat\_25062003.163648.5147/app.query.fasta.1.782  
-DB=genembl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=plc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10075846.ecgn.1.1.384.4.runat.25062003.163648.5147 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGESUBSTR -NEG\_SCORES=0 -WAIT -DSPBLOCK=200 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=6 -ICPU=6 -Fgapop=6  
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -DELOP=6 -DELEX=7

Database :  
1: genEmbl:\*  
2: gb\_ba:\*  
3: gb\_hkg:\*  
4: gb\_in:\*  
5: gb\_om:\*  
6: gb\_ov:\*  
7: gb\_pat:\*  
8: gb\_ph:\*  
9: gb\_pl:\*  
10: gb\_pt:\*  
11: gb\_ro:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_da:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_ju:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sy:\*  
28: em\_un:\*

29: em\_vl:\*  
30: em\_hkg\_hum:\*  
31: em\_hkg\_inu:\*  
32: em\_hkg\_other:\*  
33: em\_hkg\_mus:\*  
34: em\_hkg\_pln:\*  
35: em\_hkg\_fod:\*  
36: em\_hkg\_mam:\*  
37: em\_hkg\_vrt:\*  
38: em\_sy:\*  
39: em\_hkg\_hum:\*  
40: em\_hkg\_mus:\*  
41: em\_hkg\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2163	96.0	1874	6 AX392950	AX392950 Sequence
2	2157	95.8	1600	6 AX037565	AX037565 Sequence
3	1886	83.7	1251	1 AF462147	AF462147 Mus muscu
4	1749.5	77.7	2045	5 DRE404970	AJ404970 Dantio rer
5	1682.5	74.7	1857	9 HSGLYRA1	X57008 H.sapiens a
6	1677	74.5	1359	10 RNO310837	AJ310837 Rattus no
7	1677	74.5	3160	10 RNIGRAA2	X61159 R.norvegicu
8	1677	74.5	3865	10 RNNEGLY	X57281 Rat NGLYR m
9	1609	71.4	3069	9 HSU93917	U93917 Human glyci
10	1598	71.0	1443	10 RNO310838	AJ310838 Rattus no
11	1592	70.7	2413	10 RATIGRAS	M55250 Rat inhibit
12	1580.5	70.2	1715	9 HSGLYRA2	X52009 H.sapiens a
13	1580	70.2	1744	4 AF268375	AF268375 Bos tauru
14	1578.5	70.1	1392	10 AF362764	AF362764 Mus muscu
15	1577	70.0	1628	5 DRE5812	AJ005812 Dantio rer
16	1570.5	69.7	1407	10 S73717	S73717 GIRAI+Inhib
17	1569.5	69.7	1350	10 RNO310834	AJ310834 Rattus no
18	1569.5	69.7	1431	10 S73718	S73718 Girai+Inhib
19	1569.5	69.7	2125	10 RATGLYRA1	DO0833 Rattus norv
20	1568.5	69.6	1374	10 RNO310835	AJ310835 Rattus no
21	1568.5	69.6	1679	10 RNIGRA1	X55246 R.norvegicu
22	1564.5	69.5	1384	10 RNRGRTR	V00276 R.norvegicu
23	1559.5	69.2	1682	5 AF488379	AF488379 Dantio rer
24	1554	69.0	1418	5 AF094574	AF094574 Morone am
25	1537.5	68.3	2376	5 DRE308516	AJ094575 Morone am
26	1494.5	66.4	3085	5 AY094975	AY094975 Morone am
27	1413	62.7	1222	10 RNO310836	AJ310836 Rattus no
28	1381	61.3	1595	5 DRE308517	AJ308517 Dantio rer
29	983.5	43.7	657	6 E03608	E03608 DNA encodin
30	876.5	38.9	2160	5 DRE404971	AJ404971 Dantio rer
31	871	38.7	2404	6 AX305334	AX305334 Sequence
32	871	38.7	2404	6 MU093939	MU093939 Mus musculu
33	863.5	38.3	1491	10 RNO310839	AJ310839 Rattus no
34	858.5	38.1	1465	10 MMGRBMA	X81202 M.musculu
35	855.5	38.0	1737	9 AF094754	AF094754 Homo sapi
36	855.5	38.0	1737	9 AF094755	AF094755 Homo sapi
37	855.5	38.0	2106	9 H5033267	U33267 Human glyci
38	855.5	38.0	2649	9 BC032635	BC032635 Homo sapi
39	850.5	37.8	1661	4 AF268376	AF268376 Bos tauru
40	761	33.8	39796	6 AX037570	AX037570 Sequence
41	761	33.8	39796	6 H503563	Z93848 Human DNA s
42	757	33.6	1446	6 AX477445	AX477445 Sequence
43	757	33.6	1446	6 AX477446	AX477446 Sequence
44	757	33.6	5503	6 AX477442	AX477442 Sequence
45	757	33.6	5503	6 AX477444	AX477444 Sequence

RESULT 1

## ALIGNMENTS



[illegible]

QY	181	MelaspilielglnThrCysThmeGlnleuGlnSerSerIleLeuCysSerProleu	200
Db	826	ATGGACATCCAGAGCTGCAGCATCGAGTGGAG-----	858
QY	201	ProSerleuSerleuSerValGlyTyrThrmetLysAspleuValPheGluTyrPleuGlu	220
Db	859	-----AGCTTTGGCTTACCCATGAAACCCCTCGTTGGATGGCTGGAA	903
QY	221	AspAlaProAlaValGlnValAlaGlnGlyleuThrLeuProGlnPheIleLeuArgAsp	240
Db	904	GATGCTCTGCTGTGCCAAGTGGCTGAGGGGCTGACTCTGCCCAAGTTATCTTGGGGAT	963
QY	241	GluLysAspleuGlyCysCysThrLysHisTyrAsnThrGlyLysPheThrCysIleGlu	260
Db	964	GAGAAAGATCTAGAGCTGTTGTACCAACAGACATACAACACAGGAAATTCACCTGCATGAG	1023
QY	261	ValLysPheHisLeuGlnArgGlnmeGlyTyrTyrleuIleGlnmetTyrIleProSer	280
Db	1024	GTAAGTTTCACTCGGAACGGCAGATGGGCTACTATCTGATTCAAGATGCATCCAGC	1083
QY	281	LeuLeuIleValIleLeuSerTyrPValSerPheThrPileasmeAspAlaAlaProAla	300
Db	1084	CTACGCATCGATCATCTGTCTCCGTGGGTCCCTTCGAGATCAACATGATGCTGCTGCC	1143
QY	301	ArgValIleLeuGlyIleThrThrValLeuThrmetThrThrGlnSerSerGlySerArg	320
Db	1144	CGTGGGGCCCTGGGGCATCACACCGCTGCACATGACACACCGAGCTGGCTCCCGG	1203
QY	321	AlaSerLeuProLysValSerTyrValLysAlaIleAspIleTyrPmetAlaValCysLeu	340
Db	1204	GCCTCTTGCTTAAGGTCTCTACTGTAAGGCATCGACATCTGATGGCTGTGTCTGTG	1263
QY	341	LeuPheValPheAlaAlaLeuLeuGlnTyrIleAlaAlaIleAsnPheValSerArgGlnHis	360
Db	1264	CTCTTTGTGTTGCGTCCCTTCTCGAGATATGCTGCATAAATTTGTTTCTCGTCACAGAT	1323
QY	361	LysGluPheIleArgLeuValArgArgGlnArgGlnArgGlnArgGlnGluGlnAspIleIle	380
Db	1324	AAAGATTTCATTCGCTTCGAAAGAGGAGGAGGCGCCCAACCTTGGAGGAAGATATCTC	1383
QY	381	GlnGlnSerArgPheTyrPheArgGlyTyrGlyLeuGlnHisCysLeuGlnAlaIleArgAsp	400
Db	1384	CAGAAAGTCGTTCTATTTCGCTGATGGCTTGGGCGCATGCGCTGCAGGCAAGAGAT	1443
QY	401	GlyGlyProMetGlnGlySerGlyIleTyrSerProGlnProProAlaProLeuLeuArg	420
Db	1444	GGAGGCTCAATGAGAGGTTCTGGCACTTATATGTGTCCCAACCTCCAGCCCTTCTTAAGG	1503
QY	421	GluGlyGlnThrThrArgLysLeuTyrValAsp	431
Db	1504	GAAAGAGAAACACCGCGGAACCTCTAGCTGGAC	1536
RESULT 2			
LOCUS	AX037565	1600 bp	DNA
DEFINITION	Sequence 1 from Patent WO0058461.		linear
ACCESSION	AX037565		
VERSION	AX037565.1		
KEYWORDS	GI:11226986		
SOURCE			
ORGANISM	synthetic construct.		
REFERENCE	synthetic construct		
AUTHORS	artificial sequences.		
TITLE	1 (bases 1 to 1600)		
JOURNAL	Rappold-Hoebrand, G.		
	Gene for ataxia		
	Patent: WO 0058461-A 1 05-OCT-2000;		
	RAPPOLD HOERBRAND GUDRUN (DE)		
FEATURES	Location/Qualifiers		
source	1..1600		
	/organism="synthetic construct"		
	/db_xref="taxon:32630"		
	/note="CDNA"		
	1..1254		



CDS  
1. .1251  
/gene="Glyr4"  
/note="ligand-gated ion channel"  
/codon\_start=1  
/product="glycine receptor alpha 4 subunit"  
/protein\_id="AA169899.1"  
/db\_xref="GI:18448711"  
/translation="MSPDPLDKLMDRTSGYDARIKPNKGPVNVTCNIFINSEGSV  
TETMDYRVNVLKQNDPRLAYREYDDSLDNLNPSLESIMKPDLEFANEKANFH  
EETDNDKLRIFKNGNVLSTRLILISCPMDLKNFMDIOTCMOLESPGYTNDLM  
FEMLEDAPOVAVAGLTIPORILRDEKLGVTCTHYNMGKFTCEVKRLEROMGYL  
IOMTPELILYILSNVSRFIMMDAPARVAGITTVLMTQSSSRASLEKSYVKA  
IDIMAVCLLEFVALLLEYAAVNFVSRHKEFMRRLRRORRMEEDIRSRFTFRG  
YLGHCLOARDGPMGESSIYSPQPPPLKEGETMRKLYVDRAKRIDTISRVPFT  
FLVNFIFWVYKVLKSEDIHQAL"  
BASE COUNT 324 a 314 c 281 g 332 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 6.05e-182 Length: 1251  
Score: 1886.00 Matches: 358  
Percent Similarity: 94.36% Conservative: 10  
Best Local Similarity: 91.79% Mismatches: 8  
Query Match: 83.75% Indels: 14  
DB: 10 Gaps: 1  
US-10-075-846-4 (1-431) x AF462147 (1-1251)  
OY 42 MetSerProSerAspPheLeuAspLysLeuMetGlyArgThrSerGlyTyrAspAlaArg 61  
DB 1 ATGTCCCTCTGATTTCTTGACAAAGCTTATGGACACACATCTGGATATCATCTTAG 60  
OY 62 ILAARGProAsnPhelYsglyProProValAsnValThrCysAsnIlePheIleAsnSer 81  
DB 61 ATTAGGCCCAATTCAAAGGTCACCTGATGTGACCTGCACATCATCTTCAACAGT 120  
OY 82 PheSerSerValThrLysThrThrMetAspTyrArgValAsnValPheLeuArgLgn 101  
DB 121 TTTGGCTCGCTACGAGACCCACCATGGACCTACGGGTAATGTCTTCTGCGGACGCG 180  
OY 102 TTPAsnAspProArgLeuSerTyrArgLgnTyrProAspSerSerLeuAspLeuAspPro 121  
DB 181 TGGATGAGCCACGCTGCTGCGCCTACGAGATATCCAGATGACTCTGACCTAAACCA 240  
OY 122 SerMetLeuAspSerIleTyrLysProAspLeuPhePheAlaAsnGlyLysAlaAsn 141  
DB 241 TCCATGCTCGATCTATCTGGAACACGACCTCTTCTTGCACATGAGAAAGGCGCAAC 300  
OY 142 PheNlsGluValThrThrAspAsnLysLeuLeuArgLlePheLysAsnGlyAsnValLeu 161  
DB 301 TTCCATGAGGTGACCAACACAAAGTTACTGCGCATCTTCAAGAATGGAAATGTTCTC 360  
OY 162 TyrSerIleArgLeuThrLeuIleLeuSerCysLeuMetAspLeuLysAsnPhePromet 181  
DB 361 TACAGATGAGATGAGCTCATTTTGTCTGCCCAATGAGACTCAAGAACTTCCCAATG 420  
OY 182 AspIleGlnThrCysThrMetClnLeuGlnSerSerIleLeuCysSerProLeuPro 201  
DB 421 GACATACAGACCTGTACATGACGCTGGA----- 450  
OY 202 SerLeuSerLeuSerValGlyTyrThrMetLysAspLeuValPheGlnTyrLeuGluAsp 221  
DB 451 -----AGCTTTGGCTACACCAATGATGACCTATGTTTAGTGCTAGAGAGAT 498  
OY 222 AlaProAlaValGlnValAlaGlnGlyLeuThrLeuProGlnPheIleLeuArgAspGlu 241  
DB 499 GGTCCGCTGCTGCAAGTGTGAGGGGTGAGCTGACCTACCCCAATTTATTTGCGGATGAG 558  
OY 242 LysAspLeuGlyCysCysThrLysHisTyrAsnThrGlyLysPheThrCysIleGluVal 261  
DB 559 AAGGATCTAGGGGATATGTACCAAGCACTCAATACAGGAAATAATTCATTGCATTGAGGTA 618  
OY 262 LysPheNlsLeuGlnGlnMetGlyTyrTyrLeuIleGlnMetTyrIleProSerLeu 281

DB 619 AAGTTTACCCTGGAGCGGACATGGGCTACTGATCTGATTCAAATGATCATCCCAAGCTTA 678  
OY 282 LeuIleValIleLeuSerTyrPheValSerPheThrPheIleAsnMetAspAlaAlaProAlaArg 301  
DB 679 CTCATGCTGATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 738  
OY 302 ValGlyLeuGlyIleThrThrValLeuThrThrThrThrThrThrThrThrThrThrThrThr 321  
DB 739 GTAGGCTTGGCATCACACACAGTCTCTCCATGACATGCAATCTGAGACTCTGGATCCAGGGCC 798  
OY 322 SerLeuProLysValSerTyrValIleValAlaIleAspIleThrPheMetAlaValCysLeuLeu 341  
DB 799 TCTTGGCTTAAGGTGCTCTACGATTAAGGCAATTCATCTGATGATGCTGATGCTGCTGCTGCT 858  
OY 342 PheValPheAlaAlaLeuLeuGlnTyrAlaAlaIleAsnPheValSerArgLnisLys 361  
DB 859 TTTGTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 918  
OY 362 GluPheIleArgLeuArgArgArgGlnArgGlnArgGlnArgGlnArgGlnArgGlnArgGln 381  
DB 919 GAATTTATGAGACTTCTGTAAGGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 978  
OY 382 GluSerArgPheTyrPheArgGlyTyrGlyLeuGlyHisCysLeuGlnAlaArgAspGly 401  
DB 979 GAGAGCCGCTTCTATTTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1038  
OY 402 GlyProMetGluGlySerGlyIleTyrSerProGlnProProAlaProLeuLeuArgGlu 421  
DB 1039 GGTCCATGGAAGGTTCCAGCATTTATAGTCCCAACCTCCAAACCCCTTCTTAAGGAA 1098  
OY 422 GlyGluThrThrArgLysLeuTyrValAsp 431  
DB 1099 GGAGAACCATGAGGAAACTGATGTGAGC 1128  
RESULT 4  
DRE404970 2045 bp mRNA linear VRT 21-NOV-2000  
LOCUS Danto rerio mRNA for glycine receptor alpha2 subunit (glyr alpha  
DEFINITION  
ACCESSION AJ404970.1 GI:11322383  
VERSION glycine receptor alpha2 subunit; glyr gene.  
KEYWORDS zebrafish.  
SOURCE  
ORGANISM Danto rerio  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;  
Cypriniformes; Cyprinidae; Danio.  
REFERENCE 1 (bases 1 to 2045)  
Imboden, M., de Saint Jan, D., Leulier, F., Korn, H., Goblet, C. and  
Bregestovski, P.  
Isolation and Characterization of a alpha 2-type zebrafish Glycine  
Receptor Subunit  
2 (bases 1 to 2045)  
Goblet, C.  
Direct Submission  
Submitted (18-JUL-2000) Goblet C., Biotechnologies, Institut  
Pasteur, 25 rue du Dr Roux, 75724-Paris Cedex 15, FRANCE  
FEATURES  
source  
1. .2045  
/organism="Danto rerio"  
/db\_xref="taxon:7955"  
/issue\_type="brain"  
/dev\_stage="adult"  
1. .2045  
/gene="glyr alpha"  
1. .519  
/gene="glyr alpha"  
5' UTR  
520. .1850  
/gene="glyr alpha"  
/codon\_start=1  
/product="glycine receptor alpha2 subunit"

```
/protein_id="CAC16687.1"
/db_xref="GI:11322384"
/translation="MTFSLIMDCINCSCHYVFSGPSISLPLCFILYLMKMPSPS
DPLDKMGRTSGYDARIPRNGKPPYNNATFINSPGSETETMTXRYLNVRROOM
DPLATSEYRPSADSLIDPMSLDSIKRDLPEFANEKANPHEVETDKLNFONGVNL
YSRLTLIISCMDLKNFPMIDIQCTMOLSEFGYTNNDLFEMLSDNPOVADLPLP
OFVAKBEKDGVCYCTKHNTGKFCIEVKHLEKROMGYLLQMYISLLIVISWSEFM
INMDAPARVGLGITVLTMTTOSGSRASLPRYSVKALIDIMAVACLLIVFALLEY
AAVNFVSRHKEITRLRKORRIEDILYRESRFPYRPGYGHGLORMDKDRVAGS
SVAPPPYQVVLDEAVRKRFRDRAKRIDTISRATFPLSFLFNFTYITVYLVNHE
DIHNP"
1891..2045
/feature="glyc alpha"
BASE COUNT 524 a 501 c 460 g 560 t
ORIGIN
Alignment Scores:
Pred. No.: 8,72e-168 Length: 2045
Score: 1749.50 Matches: 334
Percent Similarity: 84.91% Conservative: 26
Best Local Similarity: 78.77% Mismatches: 25
Query Match: 77.69% Indels: 39
DB: Gaps: 4
US-10-075-846-4 (1-431) x DREA04970 (1-2045)
Oy 9 LeuSerPheLeuLeuLeuTrpThrLeuProGlyGlnValLeuLeuArgValAlaLeuAla 28
Db 610 CTTGCTTTCTCTATCTCTG----- 630
Oy 29 LysGluGluValLysSerGlyThrLysGlySerGlnProMetSerProSerAspPheLeu 48
Db 631 -----CAGAGCCTATGTCCTCCATGGACTTTCG 660
Oy 49 AspLysLeuMetLysArgThrSerGlyTyrAspAlaArgLeuProAsnPhelyGly 68
Db 661 GACAACTGATGGAGGAGAACCTCGGTATGATGTCGATCAGACCACTTCAAGGA 720
Oy 69 ProProValAsnValThrCysAsnIlePheIleAsnSerPheSerValThrLysThr 88
Db 721 CCACCGTAACCTCAGCTCTTACATCTTATCAGACGTTTGGATCTATTAAGACA 760
Oy 89 ThrMetAspTyrArgValAsnValPheLeuArgGlnGlnTrpAsnAspProArgLeuSer 108
Db 781 ACAATGGATACAGGCTAAACGCTTCTACGACAGAGAGGAGAACCTAGACGCGC 840
Oy 109 TyrArgGluTyrProAspSerSerLeuAspLeuAspProSerMetLeuAspSerIleTrp 128
Db 841 TACAGTGAATATCCCATGATCTCTAGACTGGACCTTCTATGTTGACTCCATATG 900
Oy 129 LysProAspLeuPhePheAlaAsnGluLysGlyAlaAsnPhelyGlnValThrThrAsp 148
Db 901 AAACCTGACTGTTTGTCTTAATGAGAAAGGTGCCAATTCATGAGTACACACAGAC 960
Oy 149 AsnLysLeuLeuArgIlePheLysAsnGlyAsnValLeuTyrSerIleArgLeuThrLeu 168
Db 961 AACACGCTGCTGAGATCTTCAAGATGGGATGCTTTACAGCATCAGGCTAACACCTC 1020
Oy 169 IleLeuSerCysLeuMetAspLeuLysAsnPheproMetAspIleGlnThrCysThrMet 188
Db 1021 ATCTCTCTCTGTCCTCCATGCTGAAGAAATTTCCCATGATGACCTGTACCATG 1080
Oy 189 GlnLeuGlnSerSerSerIleLeuCysSerProLeuProSerLeuSerLeuValGly 208
Db 1081 CACCTAGAA-----ACCTTTGGC 1098
Oy 209 TyrThrMetLysAspLeuValPheGluTrpLeuGlnAspAlaProAlaValGlnValAla 228
Db 1099 TACACCATGAAACATGATCTGAGTGGCTTCTGATTAACCT---GTGCAAGTTGGC 1155
Oy 229 GluGlyLeuThrLeuProGlnIleLeuArgAspGluLysAspLeuGlyCysCysThr 248
Db 1156 GATGACTTGTACTCTCTCAGTTTGTACTAAAGAGGAGAAAGATCTCGGCTACTGCACT 1215
```

```
Oy 249 LysHisTyrAsnThrGlyLysPheThrCysIleGluValLysPheHisLeuArgGln 268
Db 1216 AACCACTACACACAGAGTAATTCACCTGATGAGGTGAAGTTTCACTTAGAGACAG 1275
Oy 269 MetGlyTyrTyrLeuIleGlnMetTyrIleProSerLeuLeuIleValIleLeuSerTrp 288
Db 1276 ATGGGCTACTATCATTCATGATGATGATCATCCAGCGCTGATTTGATTTGCTGTGG 1335
Oy 289 ValSerPheTrpIleAsnMetAspAlaAlaProAlaArgValIleLeuGlyIleThrThr 308
Db 1336 GTGCTTTTCTGATCATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1395
Oy 309 ValLeuThrMetThrThrGlnSerSerGlySerArgAlaSerLeuProLysValSerTrp 328
Db 1396 GTGCTGACATGACACACACAGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1455
Oy 329 ValLysAlaIleAspIleTrpMetAlaValCysLeuLeuPheValPheAlaAlaLeuLeu 348
Db 1456 GTGACGGCCATTCATCATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1515
Oy 349 GluTyrAlaAlaIleAsnPhelySerArgGlnHisLysGluPheIleArgLeuArgArg 368
Db 1516 GACTATGACACACACACACACCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1575
Oy 369 ArgGlnArgArgGlnArgLeuGluGluAspIleIleGlnIleSerArg---PheTyrPhe 387
Db 1576 AACCAAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1635
Oy 388 ArgGlyTyrGlyLeuGlyHisCysLeuGlnAlaArgAspIleGlyProMetGlyGlySer 407
Db 1636 CGAGGGTACGAGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1695
Oy 408 GlyIleTyrSerProGlnProProAlaProLeuLeuArgGluGlyGluThrThrArgLys 427
Db 1696 AGGTATTCCTCCACACACACACCTGATGATGATGATGATGATGATGATGATGATGATGAT 1755
Oy 428 LeuTyrValAsp 431
Db 1756 CGCTTCTGTGAC 1767
RESULT 5
HSGLYRAL 1857 bp mRNA linear PRI 28-MAY-1993
LOCUS 1857
DEFINITION H.sapiens alpha-2 strychnine binding subunit of inhibitory glycine
receptor mRNA.
ACCESSION X52008.1 GI:31848
VERSION 1
KEYWORDS glycine receptor; inhibitory glycine receptor; strychnine binding.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1857)
Grenningloh,G., Schmieden,V., Schofield,P.R., Seeburg,P.H.,
Siddiqui,T., Mohandas,T.R., Becker,C.M. and Betz,H.
Alpha subunit variants of the human glycine receptor: primary
structures, functional expression and chromosomal localization of
the corresponding genes
EMBO J. 9 (3), 771-776 (1990)
JOURNAL MEDLINE 90183975
PUBMED 2155780
FEATURES
source
location/Qualifiers
1..1857
/organism="Homo sapiens"
/db_xref="taxon:9606"
/rnase_type="brain"
/dev_stage="fetal"
388..1746
/gene="GLYRA1"
388..1746
/gene="GLYRA1"
/note="strychnine binding alpha-2 subunit"
```



```

/product="glycine receptor alpha 2 precursor"
/protein_id="CAC35981.1"
/db_xref="GI:13548661"
/translation="MNRQVNLITLALFAFLGTHNREARCKDHSKHPSSQLSP
SDFLDKMGRTSGDARIKPNFKGPPVAVNCNIFINFGSVTEETTMQYRNIFLRQW
NDSRLAYSEYPPDSDLDIPSMIDSIKPDLEFANEKANPHVDITDKLRIKNGKV
LYSTRILTLSCPMIDLNKPNPDYQOTMOLESEFGYTMNDILFEVLISGPVAGLTL
POFLKEKEKGCCTKHNTGKFCIEVKFHLEROMGYLIOMYIPSLLIYISWSEF
MINMDAPARVAIGITVLTMTTOSGSRASLPKSVYKALDIIMAYCILEFPALE
YAAVNFYRQHKFELRLRRQKRNKEEDYTRERENFSGCMGHCLQVKGDIYKAT
PANLPDPKPDADAIKKKFEVDRAKRIIDITISNAAPLFLIFNIFWITYKAIIRHEDVH
KK"
BASE COUNT      379 a      312 c      306 g      362 t
ORIGIN
Alignment Scores:
Pred. No.:      1.18e-160      Length:      1359
Score:          1677.00      Matches:      327
Percent Similarity: 82.91%      Conservative: 32
Best Local Similarity: 75.52%      Mismatches: 48
Query Match:    74.47%      Indels:      26
Db:             10      Gaps:      6
US-10-075-846-4 (1-431) x RNO310837 (1-1359)
OY      1 MetThrThrLeuValProAlaThrLeuSerPheLeuLeuLeuTrpThrLeuProGlyGln 20
Db      13 CTACTGACACATTTGACAGCCTTGTTCATTTTCTTA-----GGGACA 57
OY      21 ValLeuLeuArgValAlaLeuAlaLysGlu---GluValLysSerGlyThrLysGlySer 39
Db      58 AACCATTCAGAGGAAGCATTCGCAAGACCATGATCCAGCTCGAAMAACATCCCTCG 117
OY      40 GlnProMetSerProSerAspPheLeuAspLysLeuMetGlyAaGthrSerGlyTyrAsp 59
Db      118 CAGACCTGCTCTCTGATGATTTCTTGATTAACATAATGGGAACATCGATGATGAT 177
OY      60 AlaArgLLeuArgProAsnPhelLysGlyProProValAsnValThrCysAsnIlePheIle 79
Db      178 GCAGAGATTCAGGCCAATTTTAAAGTCCTCCAGTAAACGTTACTTCGATATTTTATTC 237
OY      80 AsnSerPheSerSerValThrLysThrThrMetAspTyrArgValAsnValPheLeuArg 99
Db      238 AACGATTTGGATGGTGCACAGAAACACCATGATGATACCGAGATGATTTTCTGGA 297
OY      100 GlnGlnTrpAsnAspProArgLeuSerTyrArgGluLysProAspAspSerLeuAspLeu 119
Db      298 CAGCAGTGAACGATTCACGCGCTGGCATACAGTACCAACCATGATTCCTCGATTTG 357
OY      120 AspProSerMetLeuAspSerIleTrpLysProAspLeuPhePheAlaAsnGluLysGly 139
Db      358 GATCCATCGATCGCATTCGATTTGGAACACATTTGTTCTTTCGCAATGACAAAAGA 417
OY      140 AlaAsnPhelLysGluValThrThrAspAsnLysLeuLeuArgIlePheLysAsnGlyAsn 159
Db      418 GCCAATTTCCATGATGTCCACACATGATACCAAGTTGCTGGCATTTCCAAAATAGGCAA 477
OY      160 ValLeuLysSerIleArgLeuThrLeuIleLeuSerCysLeuMetAspLeuLysAsnPhe 179
Db      478 GTGCTTACAGATTAATTAACATCACTTATTCCTCCCATGAGCACTCAAAACATTT 537
OY      180 ProMetAspIleGlnThrCysThrMetGlnLeuGluSerSerIleLeuCysSerPro 199
Db      538 CCAATGATTCGACGACCTGACAAATGCACCTGGAG-----573
OY      200 LeuProSerLeuSerLeuValGlyTyrThrMetLysAspLeuValPheGluTrpLeu 219
Db      574 -----AGTTTGGGTACACCATGATGATGATGATTTGACTGGTTA 615
OY      220 GluAspAlaProAlaValAlaGluGlyLeuThrLeuProGlnPheIleLeuArg 239
Db      616 AGTGATGCTCA---GTACAAAGTTGTGAAGACTCACCTCGCTCAGTTATTTTGA 672

```

```

OY      240 AspGluLysAspLeuGlyCysCysThrLysHisTyrAsnThrGlyLysPheThrCysIle 259
Db      673 GAAGAGAAAGAACATTTGGCTATTCACAAAGCATTAACAACATTCGCAAGTTTACCTGCATT 732
OY      260 GluValLysPheHisLeuGluArgGlnMetGlyTyrTyrLeuIleGlnMetTyrIlePro 279
Db      733 GAGGTCAAGTTTACACCTGGAAGACGACATGGGCTATATTATTTGATTCAGATGATATCC 792
OY      280 SerLeuLeuIleValIleLeuSerTrpValSerPheTrpIleAsnMetAspAlaPro 299
Db      793 AGCTGTGTGATAGTCATTTTGTCTGCTCCCTTTGGATTAACATGATGACAGCTCT 852
OY      300 AlaArgValGlyLeuGlyIleThrThrValLeuThrMetThrTrpGlnSerSerGlySer 319
Db      853 GCCAGGTTGGCTTGGCATTCACACCGCTCTGCAATGATACCCAGATTCAGGTTCC 912
OY      320 ArgAlaSerLeuProLysValSerTyrValLysAlaIleAspIleTrpMetAlaValLys 339
Db      913 AGGCAATCTCGCCAAAGGCTCATATGTGAACCAATGTGACATCTGGATGGCAGCTGTC 972
OY      340 LeuLeuPheValPheAlaAlaLeuLeuGluTyrAlaAlaIleAsnPheValSerGln 359
Db      973 CTCTGTTTGTGTGTGCTGCTTACTGGAATATCACACAGTGAACATTTGCTCCAGCAA 1032
OY      360 HisLysGluPheIleArgLeuArgArgGlnArgGlnArgGlnArgGluAspIle 379
Db      1033 CACAAAGGTTCTTCTGCTCCGAGACGACAGAAAGGCGAATAAGAAAGAAATGTT 1092
OY      380 IleGlnLysSerArgPheTyrPheArgGlyTyrGlyLeuGlnHisCysLeuGlnAlaArg 399
Db      1093 ACTGCTGAAGCCGTTTAACTTACAGCGGTATGAGATGGGTCACTCCCTCCAGTGA 1152
OY      400 AspGlyLysProMetGluGlySerGlyIleTyrSer---ProGlnProAlaProLeu 418
Db      1153 GATGTACAGCTGCAAGGCTACACCTGCCAACCCACTTCCGCAACCCCA-----1203
OY      419 LeuArgGluGlyGluThrThrArgLysLeuTyrValAsp 431
Db      1204 ---AAAGTGCAGATGCTATCAACAAGACATTTGGAT 1239
RESULT 7
RNGRMA2      3160 bp      mRNA      linear      ROD 23-MAR-1993
LOCUS      R.norvegicus mRNA for inhibitory glycine receptor alpha 2A subunit.
DEFINITION      X61159
ACCESSION      X61159
VERSION      X61159.1 GI:288344
KEYWORDS      Inhibitory glycine receptor alpha subunit.
SOURCE      Rattus norvegicus.
ORGANISM      Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE      1 (bases 1 to 3160)
AUTHORS      Kunze,J., Kuryatov,A., Mauley,Y., Malosio,M.L., Schmieden,V. and
Beetz,H.
TITLE      Alternative splicing generates two isoforms of the alpha 2 subunit
of the inhibitory glycine receptor
JOURNAL      FEBS Lett. 283 (1), 73-77 (1991)
MEDLINE      91243883
PUBMED      1645300
FEATURES
source
1..3160
location/Qualifiers
1..3160
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
552..1910
/codon_start=1
/product="inhibitory glycine receptor alpha 2A subunit"
/db_xref="GI:288345"
/db_xref="SWISS-PROT:P22771"
/translation="MNRQVNLITLALFAFLGTHNREARCKDHSKHPSSQLSP
SDFLDKMGRTSGDARIKPNFKGPPVAVNCNIFINFGSVTEETTMQYRNIFLRQW
NDSRLAYSEYPPDSDLDIPSMIDSIKPDLEFANEKANPHVDITDKLRIKNGKV

```



LYSIRLTLTLSCPDLNKFNPMVDVOTQWLESFGYTMNDLIFENLSDGPVQVAGLTL  
 POFILKEKEKELGYCTKHNTGKFTQIEKFEKLEKROMGYLYLQMTPLSLIYLISWSP  
 WINDADAPARVALGITTIVLTMTTQSSGSAKLPRYSYKALDIIMAVCLLEFPALAE  
 YAAVNFVSRHKEFLRLRRROKRONKEDVRESRPNRGVGMCHLOVKRGTVAKAT  
 PANLPQPRADADAIKKKFDVRAKRIDTISNAAPFLFLININITYWITYKIRIEDVH  
 KK\*

sig\_peptide 552. .632  
 mat\_peptide 630. .1907  
 /product="inhibitory glycine receptor alpha 2A subunit"

BASE COUNT 902 a 669 c 645 g 944 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 3.65e-160 Length: 3160  
 Score: 1677.00 Matches: 327  
 Percent Similarity: 82.91% Conservative: 32  
 Best Local Similarity: 75.52% Mismatches: 48  
 Query Match: 74.47% Indels: 26  
 DB: 10 Gaps: 6

US-10-075-846-4 (1-431) x RNIGRAA2 (1-3160)

QY 1 MethThrLeuValProAlaThrLeuSerPheLeuLeuLeuTrpThrLeuProGlyGln 20  
 Db 564 CTAGGAACTATTGGACAGCCTTGTTCGATTTCCTTA-----GGGACA 608  
 QY 21 ValLeuLeuArgValAlaLeuAlaGlu---GluValLysSerGlyThrLysGlySer 39  
 Db 609 AACCACTTCAGGAGGAACTCTGCNAAGACCATGCAGTCCAGTGGAAACATCCCTC 668  
 QY 40 GlnProMetSerProSerAspPheLeuAspLysLeuMetGlyArgThrSerGlyLysArg 59  
 Db 669 CAGACCCGTCTCCCTTCAGATTCTTGATTAACAAAGGAGGAGGACATCAGGATATAT 728  
 QY 60 AlaArgIleArgProAsnPhelGlyProProValAsnValThrCysAsnIlePheIle 79  
 Db 729 GCAAGAAATCAGGCCAAATTTTAAAGCTCTCCAGTAAACGTTACTTGCATATTTTATC 788  
 QY 80 AsnSerPheSerSerValThrLysThrThrMetAspTyrArgValAsnValPheLeuArg 99  
 Db 789 AACACTTTTGGATCGGTACAGAAACACCATGACGACTACCGAGTACATTTTCTGAGA 848  
 QY 100 GlnGlnTrpAsnAspProArgLeuSerTyrArgGluTyrTrpProAspSerLeuAspLeu 119  
 Db 849 CAGCGTGAACGATTCAGGCTGGCATACAGTACGCCAGATGATTCCTGGATTTG 908  
 QY 120 AspProSerMetLeuAspSerIleTyrLysProAspLeuPheAlaAsnGluLysGly 139  
 Db 909 GATCATCGATGCTGATTCATTTGGAACAGATTGTCTTGGCCATGAGAAAGA 968  
 QY 140 AlaAsnPhelAspGluValThrThrAspAsnLysLeuLeuArgIlePheLysAsnGlyAsn 159  
 Db 969 GCCAATTTTCATGATGATCACCACCATATACAAAGTTGCGCGGATTTCCAAAAATGGCAAA 1028  
 QY 160 ValLeuTyrSerIleArgLeuThrLeuLeuLeuSerCysLeuMetAspLeuLysAsnPhel 179  
 Db 1029 GTGCTTACAGATTAGACTACCTTGACTTTATCTCTGCTCCATGAGACTGAAAGACTTT 1088  
 QY 180 ProMetAspIleGlnThrCysThrMetGlnLeuGlnLeuSerSerIleLeuCysSerPro 199  
 Db 1089 CCAATGATGATTCAGACCTGTCATATGACAGTGGAG----- 1124  
 QY 200 LeuProSerLeuSerLeuSerValGlyTyrThrMetLysAspLeuValPheGluTrpLeu 219  
 Db 1125 -----AGTTTGGTACACCATCATGATGATATATTGAGTGGTTA 1166  
 QY 220 GluAspAlaProAlaValGlnValAlaGlnGluLeuThrLeuProGlnPheIleLeuArg 239  
 Db 1167 AGTGATGATGCA---GTACAAAGTCTGAGAGACTACCTCTCAGTTATTTTGA 1223  
 QY 240 AspGluLysAspLeuGlyCysCysThrLysHisTyrAsnThrGlyLysPheThrCysIle 259  
 Db 1224 GAAGAGAGAGAACTTGGCTATTGGACAAAGCATTTACACACTGGCAGTTTACCTGCAT 1283

QY 260 GluValLysPheHisLeuGluArgGlnMetGlyTyrTyrLeuIleGlnMetCysIlePro 279  
 Db 1284 GAGGTCAAGTTTCAACCTGAGCAAGCAAGATGGGCTATATTGTGATTCAGATATATCC 1343  
 QY 280 SerLeuLeuIleValIleLeuSerTyrValSerPheTrpIleAsnMetAspAlaPro 299  
 Db 1344 AGCCTGTGATGATGATTTTCTCTGGGTCTCTTTGGATTAACATGATGAGAGCTTCT 1403  
 QY 300 AlaArgValGlyLeuGlyIleThrThrValLeuThrMetThrThrGlnSerSerGlySer 319  
 Db 1404 GCCAGAGTTGCCCTTTGGCATCACACCTCTGACAAATGATACCATCCAGATTCAGTTCC 1463  
 QY 320 ArgAlaSerLeuProLysValSerTyrValLysAlaIleAspIleTrpMetAlaValCys 339  
 Db 1464 AGGCACTCTCGCCAAAGCTCATATGTAAGCAATTCATCTGATGAGAGTGGC 1523  
 QY 340 LeuLeuPheValPheAlaAlaLeuLeuGlnLysArgAlaIleAsnPhelValSerArgGln 359  
 Db 1524 CTCTGTTGTGTGGTGGCTTACTGGAATATGACAGTGAACCTTGTCTCCAGGCA 1583  
 QY 360 HisLysGluPheIleArgLeuArgArgArgGlnArgArgGlnArgLeuGluGluAspIle 379  
 Db 1584 CACAAAGAGTTCCTCTCGCTCCGAGACAGACAGAAAGCAAGAAATGAAGAGATGTT 1643  
 QY 380 IleGlnGluSerArgPheTyrPheArgGlyTyrGlyLeuGlyHisCysLeuGlnAlaArg 399  
 Db 1644 ACTCTGTGAACCGCTTTTACTTCAGCGGTATGAGATGAGTCACTGCTCCAGTGAA 1703  
 QY 400 AspGlyLysProMetGluGlySerGlyLysLeuSer---ProGlnProAlaProLeu 418  
 Db 1704 GATGTTACAGCTGTGAAGCTTACACCTCCCAACCCACTCCGCAACCCCA----- 1754  
 QY 419 LeuArgGluGlyGluThrThrArgLysLeuTyrValAsp 431  
 Db 1755 ---AAGATGCAAGATGCTATCAAGAAAGATTGTGAT 1790

RESULT 8  
 RNNEOGLY 3865 bp mRNA linear ROD 07-AUG-1991  
 LOCUS Rat NG1YR mRNA for neonatal glycine receptor.  
 DEFINITION X57281  
 ACCESSION X57281.1 GI:56743  
 VERSION 1  
 KEYWORDS glycine receptor; NG1YR gene.  
 SOURCE Rattus norvegicus.  
 ORGANISM Rattus norvegicus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 3865)  
 AUTHORS Hishinuma, F.  
 DIRECT SUBMISSION  
 JOURNAL Submitted (28-JAN-1991) F. Hishinuma, Mitsubishi Kasei Institute of Life Sciences, 11 Minamiooya, Machida-shi, Tokyo 194, Japan  
 REFERENCE 2 (bases 1 to 3865)  
 AUTHORS Akagi, H., Hirai, K. and Hishinuma, F.  
 TITLE Cloning of a glycine receptor subtype expressed in rat brain and spinal cord during a specific period of neuronal development  
 JOURNAL FEBS Lett. 281 (1-2), 160-166 (1991)  
 MEDLINE 91200276  
 PUBMED 1707830

FEATURES  
 source  
 Location/Qualifiers  
 1..3865 /organism="Rattus norvegicus"  
 /strain="Wistar"  
 /db\_xref="taxon:10116"  
 /tissue\_type="spinal cord"  
 /dev\_stage="10 days old"  
 1..3865 /gene="NG1YR"  
 1..3865 /evidence="experimental"

CDS  
1173..2531  
/gene="NGlyR"  
/codon\_start=1  
/product="neonatal type of glycine receptor"  
/protein\_id="CAA0549.1"  
/db\_xref="GI:56744"  
/translation="NNROLVNLTALFAFLGTNHFREAFCCKDHSRSGKHPQSOTLSP  
SDFELKMGRTSGDARIRPNFKGVPVNTGFIINSFGSVETETMDRYNIELROOM  
NDSRLAYSEYPDSDLSDPSMLDSITKRPDLFANENKAGHNDYTDNKKLRISKNKV  
LVSIRTLTSLSCPMDLAKNFMVDQCTMOLSESTYTMNDLIFEMLSLDPQVQAEGLTL  
PQTLKEEKELGCTAHNTGKFTCLEVFHLEROMGYLIOWITPSLIVLISWSE  
WIMMDAPARVALGITTVMITQSSGSRSALSRVYKVIDMAVCLIFVPAALLE  
YAAVNVSRQHEFLRLRRORONKEEDVTSRENVSGYGHCLQVFGDVAVKAT  
PANLPQPKDADAIKKKFDRAKRIDTISRAPLAFLEFNIFWITTKIINHEVH  
KK"

BASE COUNT 1087 a 833 c 889 g 1056 t  
ORIGIN

Alignment Scores:  
Pred. No.: 4,78e-160 Length: 3865  
Score: 1677.00 Matches: 327  
Percent Similarity: 82.91% Conservative: 32  
Best Local Similarity: 75.52% Mismatches: 48  
Query Match: 74.47% Indels: 26  
DB: 10 Gaps: 6

US-10-075-846-4 (1-431) x RNNEOGLY (1-3865)

OY 1 MetThrThrLeuValProAlaThrLeuSerPheLeuLeuLeuThrPheLeuProGlyGln 20  
Db 1185 CTAGTGAACATTTTGGACGCTTGTGTTTCTTCTTA-----GGGACA 1229

OY 21 ValLeuLeuArgValAlaLeuAlaLysGlu---GluValLysSerGlyThrLysGlySer 39  
Db 1230 AACCACTTCAGGAGACATTCCTGCAAGACCATGACCTCCAGTCTGAAACATCCCTCG 1289

OY 40 GlnProMetSerProSerAspPheLeuAspLysLeuMetGlyArgThrSerGlyTyrAsp 59  
Db 1290 CACAGCCCTGCTCCTTCAGATTTCTTGATTAACCTAATGGAGAGACATCAGGATATGAT 1349

OY 60 AlaArgLysArgProAsnPhelyGlyProProValAsnValThrCysAsnIlePheIle 79  
Db 1350 GCACAGATCAGGCGCAATTTTAAAGGTCCTCCAGTAACCTTACCTGCAATATTTTATC 1409

OY 80 AsnSerPheSerSerValThrLysThrThrMetAspTyrArgValAsnValPheLeuArg 99  
Db 1410 AACAGTTTGGATCGGTCACAGAAACCATGACACTACCGAGTGAACATTTTCTGAGA 1469

OY 100 GlnGlnTyrPasnAspProArgLeuSerTyrArgLysTyrProAspAspSerLeuAspLeu 119  
Db 1470 CACAGTGGAGACATTCACGCGCTGGCATACAGTACAGTACCATGATTCCTCGATTTG 1529

OY 120 AspProSerMetLeuAspSerIleTyrLysProAspLeuPheLeuAsnGluLysGly 139  
Db 1530 GATCCATCGAGTGTGATTCGATTTGGAACACGATTTGTTCTTCCATGAGAAAGGA 1589

OY 140 AlaAsnPheIleGluValThrThrAspAsnLysLeuLeuArgLysPheLysAsnGlyAsn 159  
Db 1590 GCCAATTTTCCATATGTCACACCATGATACCAAGTGTGCGGATTTCCAAAATGCGAAA 1649

OY 160 ValLeuTyrSerIleArgLeuThrThrLeuLeuLeuSerCysLeuMetAspLeuLysAsnPhe 179  
Db 1650 GTGCTGACAGTATTTAGACTCACCTTGACATTATCCGTCCCATGAGACCTGGAAGAACTTT 1709

OY 180 ProMetAspIleGlnThrCysThrMetGlnLeuGluSerSerSerIleLeuCysSerPro 199  
Db 1710 CCATATGATTCAGACCTGTACAAATGCCACTGAG----- 1745

OY 200 LeuProSerLeuSerLeuSerValGlyTyrThrMetLysAspLeuValPheGluTyrLeu 219  
Db 1746 -----AGTTTGGGTACACCATGATGACCTGATATTGAGTGTTA 1787

OY 220 GluAspAlaProAlaValGlnValAlaGluGlyLeuThrLeuProGlnPheIleLeuArg 239  
Db 1788 AGTGAATGATCCCA---GTACAGATTTGGCTGAAGAGCTACCCCTGGCTCAGTTATTATGAAA 1844

OY 240 AspGluLysAspLeuGlyCysCysThrLysHisTyrAsnThrLysPheThrCysIle 259  
Db 1845 GAAGAGACAGCAACCTTGCTATTCACAAACCATATACACCTGAGCAAGTTTACCTGCATT 1904

OY 260 GluValLysPheHisLeuGluValArgGlnMetGlyTyrTyrLeuLeuGlnMetCyrIlePro 279  
Db 1905 GAGCTCAGATTTACCTGACCGACCCAGATGGCTATTATTGATCCAGATGATATTCGCC 1964

OY 280 SerLeuLeuIleValIleLeuSerTyrPheSerPheThrPheLeuMetAspAlaAlaPro 299  
Db 1965 AGCCTGATGATCATATTGTTGCTCGTCTCTCTTTGGATTAACATGATGACGCTCCT 2024

OY 300 AlaArgValGlyLeuGlyIleThrThrValLeuThrMetThrThrGlnSerGlySer 319  
Db 2025 GCCAGGCTTCCCTTGCCATCACACCCCTCTGACATGACATGCCAGACTTCAGGTTCC 2084

OY 320 ArgAlaSerLeuProLysValSerTyrValLysAlaIleAspIleThrPheAlaValCys 339  
Db 2085 AGGCACTCTGTCACAAAGGTCATATGTAAGCAATTGACATCTGATGCGAGTGTGC 2144

OY 340 LeuLeuPheValPheAlaAlaLeuLeuGluTyrAlaAlaIleAsnPheValSerArgGln 359  
Db 2145 CTTCCTGTTTGTGTTGCTGCTTACTGAGATATGACAGCATGAACTTGTCTCCAGCAA 2204

OY 360 HisLysGluPheIleArgLeuArgArgArgGlnArgArgGlnArgLeuGluAspIle 379  
Db 2205 CACAGAGATGTCCTCTGCTCCGAGACGACAGACAGGCGCAATAGGAAGAAGATGTT 2264

OY 380 IleGlnIleSerArgPheTyrPheArgGlyTyrGlyLeuGlnHisCysLeuGlnAlaArg 399  
Db 2265 ACTCGTAAAGCGTTTAACTTACCTGACGGTATGAGGATGGGTACCTGCCAAGTGAAA 2324

OY 400 AspGlyLysProMetGluGlySerGlyIleTyrSer---ProGlnProProAlaProLeu 418  
Db 2325 GATGCTACACTGTGTCACAGGCTACACCTGCCACCCACTTCCGACACCCCA----- 2375

OY 419 LeuArgGluGlyLysThrThrArgLysLeuTyrValAsp 431  
Db 2376 ---AAGATGCAGATGCTATCAGAGAGAAAGTGTGTGAT 2411

RESULT 9  
LOCUS HS093917 3069 bp mRNA linear PRI 27-JUL-1998  
DEFINITION Human glycine receptor alpha 3 subunit mRNA, complete cds.  
ACCESSION U93917  
VERSION 093917.1 GI:3342235  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 3069)  
Nikolic,Z., Laube,B., Weber,R.G., Licher,P., Kioschis,P.,  
Rouska,A., Mulhardt,C. and Becker,C.M. Glna3 gene structure,  
the human glycine receptor subunit alpha3. Glna3 gene structure,  
chromosomal localization, and functional characterization of  
alternative transcripts  
J. Biol. Chem. 273 (31), 19708-19714 (1998)  
MEDLINE 98344067  
PUBMED 9677400  
2 (bases 1 to 3069)  
REFERENCE  
AUTHORS Nikolic,Z.  
TITLE Direct Submission  
SUBMITTER Submitted (17-MAR-1997) Biochemistry, University  
Erlangen-Nuremberg, Fahrstr. 17, Erlangen 91054, Germany  
FEATURES  
source  
1..3069  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

CDS  
421..1770  
/note="ligand gated ion channel"  
/codon\_start=1  
/product="glycine receptor alpha 3 subunit"  
/protein\_id="AAC39917.1"  
/db\_xref="GI:3342236"  
/translation="MAHVRHFRILVSGFYWEAALLSLVATKETSARSRSAPMSPS  
DEFDKLMGRTSGDARIRPNFKGPVAVTCNIFINSFGSLAETIMDYRNVIFLRQKN  
DPLRAYSEYDDSLDLPMSIDSLIMKPDLEFANEGANFEDNDKRLNFRKGNV  
YSIRLTLSCEPMDLKNFMDVOTICIMOLEFSGYTMDLFEEMDEAPVOYAEGLTP  
OILKEKDLRYCTKHYNTGKFCIEVRFLROMGYLLIOMYPSLIIYLVSVSRF  
INMDAARARVAGITVLTMTTOSGSRASIPKSYKADIDMAVCLILFYSLFV  
AANFVSROHLELFRKRKNKDEVRSESTAYAGMGPCLDAGKMTPKGNHLY  
QVMKPSDEMKRKFIDRAKKIDITISRAFCPLAFLIFNIYFWVYIKILRHEDHQOD"  
BASE COUNT 962 a 547 c 583 g 977 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 2,866-153 Length: 3069  
Score: 1609.00 Matches: 310  
Percent Similarity: 83.94% Conservative: 35  
Best Local Similarity: 75.43% Mismatches: 42  
Query Match: 71.45% Indels: 24  
DB: 9 Gaps: 5  
US-10-075-846-4 (1-431) x HS093917 (1-3069)  
Oy 22 LeuLeuArgValAlaLeuAlaLysGluGluValLysSerGlyThrLysGlySerGlnPro 41  
Db 481 CTGTACACCACTTTGGTTCACAAAGAAACAGACAGTGCAGATCTCGAAGTCTCCA 540  
Oy 42 MetSerProSerAspPheLeuAspLysLeuMetGlyArgThrSerGlyTyrAspAlaArg 61  
Db 541 ATGTCACCTCTGATTTCTGATTAATTAATGAGCAGACATCATGATATGACCAAGA 600  
Oy 62 IleArgProAsnPhelGlyProProValAsnValThrCysAsnIlePheIleAsnSer 81  
Db 601 ATCAGACCCCAATTTAAAGCCCTCCAGTATATGCAATCAACATATTCATCAACGT 660  
Oy 82 PheSerSerValThrLysThrMetAspTyrArgValAsnValPheLeuArgGlnGln 101  
Db 661 TTGCGCTCTATCGAGAGACACATGATGATACAGAGATATCTTCTTCGACAGAA 720  
Oy 102 TrpAsnAspProArgLeuSerTyrArgGluTyrProAspAspSerLeuAspLeuAspPro 121  
Db 721 TGGATGATGCCCGCCCTCGTACAGTGAATATCTCTGACGACTTTAGACCTGACCCC 780  
Oy 122 SerMetLeuAspSerIleTyrLysProAspLeuPhePheIleAsnGluLysGlyValAsn 141  
Db 781 TCCATGTTGACCTCATTTGGAAACCTGATTTGTTTGGCAATGAAGGAGTCCAC 840  
Oy 142 PheHisGluValThrThrAspAsnLysLeuLeuArgIlePheLysAsnGlyAsnValLeu 161  
Db 841 TTTATGAGACGACACAGACAAATGCTAGAAATTTCAAAAAAGGAATGTTCTT 900  
Oy 162 TyrSerIleArgLeuThrLeuIleLeuSerCysLeuMetAspLeuLysAsnPheProMet 181  
Db 901 TATTCATTAAGATTAACATTAACCTTCTGCTCAATGAGATTCACAAATTTTCCCATG 960  
Oy 182 AspIleGlnThrCysThrMetGlnLeuGlnSerSerIleLeuCysSerProLeuPro 201  
Db 961 GATGACACAAACATGATATGCAACTGAA----- 990  
Oy 202 SerLeuSerLeuSerValGlyTyrThrMetLysAspLeuValPheGlnTyrPheGlnLysP 221  
Db 991 -----ACCTTGGGTACACATGATGATGATCATTTTGGATGGCAGACATAG 1038  
Oy 222 AlaProAlaValGlnValAlaGluGlyLeuThrLeuProGlnPheIleLeuArgAspGlu 241  
Db 1039 GCACCC---GTACAGAGTGGCAGAAAGACTCTTGGCCCGACTTCTGTGAAGAA 1095  
Oy 242 LysAspLeuGlyCysCysThrLysHisTyrAsnThrGlyLysPheThrCysIleGluVal 261  
|||||

Db 1096 AAGATTACGATCTGCACTAAACATTACAAATACAGAAAGTTTACGTATAGAGTG 1155  
Oy 262 LysPheHisLeuGluArgGlnMetGlyTyrTyrLeuIleGlnMetTyrIleProSerLeu 281  
Db 1156 CGATTCCATCTGGAGCGCAAAATGGATACTATCTGATCCAGATGATCCAGTCTC 1215  
Oy 282 LeuIleValIleLeuSerTyrPvalSerPheThrIleAsnMetAspAlaIleProAlaArg 301  
Db 1216 CTGATTGTTATTTATCTCTGAGTCTCTCTGATCAACATGATGACGACCGCCAGG 1275  
Oy 302 ValGlyLeuGlyIleThrThrValLeuThrMetThrThrCysLeuSerGlySerArgAla 321  
Db 1276 GTTACTCTGGGGATACCGACGCTGCTACGATGATGACACAGATTCAGATCAGAGCT 1335  
Oy 322 SerLeuProLysValSerTyrValLysAlaIleAspIleThrPheMetAlaValCysLeuLeu 341  
Db 1336 TCCTTGCCAAAGATTTCATATGTCAAAGCATATGATATTTGGATGGCAGTATGCTCTT 1395  
Oy 342 PheValPheAlaAlaLeuLeuGluTyrAlaAlaIleAsnPheValSerArgGlnHisLys 361  
Db 1396 TTGTGTTTCACGACATTCGAGTACGATGACGCTGTAATTTTGTATCAAGACAAACAAA 1455  
Oy 362 GluPheIleArgLeuArgArgArgGlnArgArgGlnArgLeuGluGluAspIleIleGln 381  
Db 1456 GAACTTCTGAGATTTGCAACGAAAGAAAGAAAT-----AAGCATGATGACGTAAAG 1506  
Oy 382 GluSerArgPheTyrPheArgGlyTyrGlyLeuGlnLysCysLeuGlnAlaArgAspGly 401  
Db 1507 GAAAGCCGATTCACCTTCACAGCCTATGAGATGGAGACATGTCTACAGCAAGAGATGCC 1566  
Oy 402 GlyProMetGlyGlySerGlyIleTyrSerProGlnProAlaPheLeuLeu--Arg 420  
Db 1567 ATGACTCCAAAGGAGC-----CCCAACCAACCTCTGCGCAATGACCAAAA 1611  
Oy 421 GluGlyGluThrThrArgLysLeuTyrValAsp 431  
Db 1612 AGTCTGATGAATGAGAGAGAGTCTTATTCGAC 1644  
RESULT 10  
RNO310838 1443 bp mRNA linear ROD 03-APR-2001  
LOCUS Rattus norvegicus mRNA for glycine receptor alpha.3 precursor.  
DEFINITION A31310838  
ACCESSION A31310838  
VERSION A31310838.1 GI:13548662  
KEYWORDS glycine receptor alpha 3 precursor.  
SOURCE Norway rat.  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 1443)  
Beato,M., Groot-Kormelink,P.J., Colquhoun,D. and Sivilotti,L.G.  
Concentration dependence of single channel currents through rat  
recombinant alpha 1 glycine receptors  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1443)  
Groot-Kormelink,P.J.  
Direct Submission  
TITLE Submitted (30-MAR-2001) Groot-Kormelink P.J., Department of  
Pharmacology, The School of Pharmacy, 29/39, Brunswick Square,  
London, WC1N 1AX, UNITED KINGDOM  
FEATURES  
source  
1..1443  
/organism="Rattus norvegicus"  
/strain="Sprague-Dawley males"  
/db\_xref="taxon:10116"  
/db\_xref="taxon:10116"  
/tissue\_type="whole brain"  
/clone\_lib="rat brain marathon-ready cDNA"  
/dev\_stage="8-12 weeks"  
1..1443  
/function="ion channel"  
/codon\_start=1  
/evidence=experimental

```

/product="glycine receptor alpha 3 precursor"
/protein_id="CAC35982.1"
/db_xref="GI:13548663"
/translation="MPWIRLFSSGSIIFRIMAHVHRRTILSGFWEALLSLVAT
KENSARSRSRSPSPDLKLFKSRISGYARIRPNKGPVNTCITFINRESSTAE
ITMDYRNIFLRKMDNDRLAYSEYDSDLDLSDLSIMKPDLEFANKGANFHEV
ITDNKRLIRKNGVNLISRLITLSCPMKNEPMQVOCIMOLEFSGYIMDLIE
MODEAPYVAEGELTLPOLKEEDLRCYKHYTGKTCIEVFEHLEROMGYLIDM
YIPSLIVLSWVSFNIMDAAPARVALGTTVLTMTTOSGSRSLPKVSYKVAIDM
MAVCLIFVSALLEVAANVAVNSROHKEITLFRPKRKNTCAEAFLEKFRSPDDDEY
RESRSTFAKGMKPCLOAKDGVYPKGPNHNVQVMPKSADEMRKYIDRAKKIDITISA
CFPLAFLEIFWVITKILRHEDIHQD"
BASE COUNT      405 a      320 c      324 g      394 t
ORIGIN

Alignment Scores:
Pred. No.:      1 37e-152      Length:      1443
Score:          1596.00      Matches:      314
Percent Similarity: 79.27%      Conservative: 34
Best Local Similarity: 71.53%      Mismatches: 49
Query Match:      70.96%      Indels:      42
DB:              10      Gaps:      7

US-10-075-846-4 (1-431) x RNO310838 (1-1443)

QY      9 LeuSerPheLeuLeuLeuTriPThrLeuProGlyGlnValLeuLeuArgValAlaLeuAla 28
Db      CTTTCTCGATTATTACTCTCGGAA-----GCCGCACGTGTACGAGCTGGTGCC 129
QY      29 LysGlnGluValLysSerGlyThrLysGlySerGlnProMetSerProSerAspPheLeu 48
Db      130 ACGAAGCAACAAACACGCGTCCCGAAGTCCCAAGTCACTTCGATTTCTG 189
QY      49 AspLysLeuMetGlyArgThrSerGlyTyrAspAlaArgLeuArgProAspPheLysGly 68
Db      190 GACAACTAATGGGAGAGACATCCGGTAGATGATCAAGAACAGCCCAACTTAAAGT 249
QY      69 ProProValAsnValThrCysAsnIlePheIleAsnSerPheSerValThrLysThr 88
Db      250 CCTCCAGCTTAATGTCACATGCAACATATTCATAACAGCTTGGCTCCATCCGACAGACG 309
QY      89 ThrMetAspTyrArgValAsnValPheLeuArgGlnIntrPAsnAspProArgLeuSer 108
Db      310 ACATATGATTACAGAGTAACATTTCTTCCTCGAGAGTGAATGATCTCCGCTTGA 369
QY      109 TyrArgGlnTyrProAspAspSerLeuAspLeuAspProSerMetLeuAspSerIleTyr 128
Db      370 TACAGTGAATACCTGATGATTCATTAGACCTCGAACCATTCATGTTGGACTCCATATGG 429
QY      129 LysProAspLeuPhePheAlaAsnGlnLysGlyAlaAsnPheHisGluValThrThrAsp 148
Db      430 AAACCTGACTTGTCTTGTGTAATGAGAAAGGGGCTCACTCCAGAACTGACACCGAT 489
QY      149 AsnLysLeuLeuArgIlePheLysAsnGlyAsnValLeuTyrSerIleArgLeuThrLeu 168
Db      490 AACCAAGCTGCTAAATAATTTTCAAAATGGAATGCTTTTAATTAATAAGGCTACATTA 549
QY      169 IleLeuSerCysLeuMetAspLeuLysAsnPheProMetAspIleGlnThrCysThrMet 188
Db      550 ACACTCTCTTGTCATATGATCTCAAGAAATTTCCCAATGGATGTTCAAACTGATTAATG 609
QY      189 GlnLeuGlnSerSerIleLeuCysSerProLeuProSerLeuSerLeuSerValGly 208
Db      610 CAACTCGAA-----AGCTTTGGG 627
QY      209 TyrThrMetLysAspLeuValPheGlnTyrPleuGlnLysAlaPheAlaValGlnValAla 228
Db      628 TACACGATGATGATCTCATTTTGAATGGCAAGATGAAAGCACA---GTACAAGTGGCT 684
QY      229 GlnGlyLeuThrLeuProGlnPheIleLeuArgAspGlnLysAspLeuGlyCysCysThr 248
Db      685 GAAGGACTCACTTGGCTCAATTTCTGTTGAAGAAGAAAGATTGGCGACTACTGACT 744

```

```

QY      249 LysHisTyrAsnThrGlyLysPheThrCysIleGlnValLysPheHisLeuGluArgGln 268
Db      745 AAACACTACATACAGGAAGTTTACATGCATAGACATGATTTCAATCTTGACGGCAA 804
QY      269 MetGlyTyrTyrLeuIleGlnMetTyrIleProSerLeuLeuIleValIleLeuSerTyr 288
Db      805 ATGGGCTACTACTTGATTCAGATGATGATCCAGCTTCCTTGATGTCATCTCTCCGCG 864
QY      289 ValSerPheThrPheLeuMetAspAlaAlaPheAlaArgValGlyLeuGlyIleThrThr 308
Db      865 GTCTCATTTCTGATTAACATGATGATGACCTCCGCTGGGATGCGGTATACCACT 924
QY      309 ValLeuThrMetThrThrGlnSerSerGlySerArgAlaSerLeuProLysValSerTyr 328
Db      925 GTACTTAGATGACACGACGACAGTTTGATGATCCGGGCTTCTTACCAAGGTGCTCAT 984
QY      329 ValLysAlaIleAspIleTyrPheMetAlaValCysLeuLeuPheValPheAlaIleLeu 348
Db      985 GTCAAGGCAATTTGACATTTGATGCGAGTGTCTCTTTTGTCTTCAGCACTTCTG 1044
QY      349 GlnTyrAlaAlaIleAsnPheValSerArgGlnHisLysGlnPheIleArgLeuArgArg 368
Db      1045 GAGTATGACGCCCGTGAAATTTTGTATCAAGGCAACAAAGACTGTGAGGTTTCGGCGA 1104
QY      369 ArgGlnArgArgGln-----ArgLeuGlu----- 376
Db      1105 AAGAGAAATAATTAACAGAAAGTTTGCATCGAGAGAAAGTTTACCGTTCTCAGACAG 1164
QY      377 GlnAspIleIleGlnGlnSerArgPheTyrPheArgGlyTyrGlyLeuGlyHisCysLeu 396
Db      1165 GATGATGAGGAGGAGGAGAGCGGTTCAGCTTCAGTCCGTCTTGAAATGGGCCCTGTCTT 1224
QY      397 GlnIleArgAspGly-----GlyProMetGlnGlySerGlyIleTyrSerPro 412
Db      1225 CAACCAAGATGATGTTGTTGTTCCAAAGGTTCCAAACCATGCTGTCCAGGTC----- 1275
QY      413 GlnProAlaProLeuArgGlnGlyGlnThrThrArgLysLeuTyrValAsp 431
Db      1276 -----ATGCCAAAGAGCGCCGATGAATGAGAGAAAGTCTTCATCGAC 1317

RESULT 11
RATIGRAS
LOCUS      2413 bp      mRNA      linear      ROD 07-MAR-1995
DEFINITION Rat inhibitory glycine receptor alpha subunit mRNA, complete cds.
ACCESSION M55250.1
VERSION M55250.1 GI:204882
KEYWORDS inhibitory glycine receptor.
SOURCE Rat (strain Wistar) adult brain, cDNA to mRNA, clone pGR48-alpha-3.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 2413)
AUTHORS Kuhnse,J., Schmidem,V. and Betz,H.
TITLE Identification and functional expression of a novel ligand binding
subunit of the inhibitory glycine receptor
JOURNAL J. Biol. Chem. 265 (36), 22317-22320 (1990)
MEDLINE 91093073
PUBMED 2176214
FEATURES
SOURCE
location/Qualifiers
1..2413
/organism="Rattus norvegicus"
/strain="Wistar"
/db_xref="taxon:10116"
/clone="pGR48-alpha-3"
/tissue_type="brain"
/dev_stage="adult"
1..2413
/gene="inhibitory glycine receptor"
400..1794
/gene="inhibitory glycine receptor"
/codon_start=1
/product="inhibitory glycine receptor alpha subunit"

```

```

/protein_id="AAAG3492.1"
/db_xref="GI:204883"
/translation="MAHVRERLTLSGFYFMEALALSLVATKETNSARSAPMS
DPLDKLMGRTSGDARIRBNPKGPVAVTCNFINFSNGSLAETMDRVNIFROKN
DPRAYSEYDDSDLDPSMLDSIMKPDLEFANEGANFHEVTTDKNLIFKGNVIL
YSTRLLTSCPMOLKNEPMDVOICIMQLSFGYTKMDLFEWODEAPVOAEGTLPL
QFLKEKRDLYCTKHNKTFCTICIEVRFLHROMGYLTOMYPSLLIILSVSWM
INMDAPARVALGTTVLTMTTOSGSGRSLPKSVYKVIDIMAVCOLLEFVALLEY
AAVNFVSRKELRFRKRNKTEAFELSEYFSDDEVRSLSEFTAYGMPCL
QAKGVYKGPBNHAYVMPKPSADEMRKVFIDRAKIDTISACPLFLFENFYWVI
YKILRHEDIHQDD"
400..498
/gene="inhibitory glycine receptor"
/mat_peptide
496..1791
/gene="inhibitory glycine receptor"
/product="inhibitory glycine receptor alpha subunit"
BASE COUNT 697 a 503 c 557 g 656 t
ORIGIN
Alignment Scores:
Pred. No.: 1,11e-151 Length: 2413
Score: 1592.00 Matches: 313
Percent Similarity: 79.04% Conservative: 34
Best Local Similarity: 71.30% Mismatches: 50
Query Match: 70.69% Indels: 42
DB: 10 Gaps: 7
US-10-075-846-4 (1-431) x RATIGRAS (1-2413)
OY 9 LeuserPheLeuLeuLeuTrPheLeuProGlyGlnValLeuLeuArgValAlaLeuAla 28
Db 430 CTTTCTGATTTTACTTCTGGGAA-----GCCGCCGCTGTACTAGTTGGTGCC 480
OY 29 LysGluGluValLysSerGlyThrLysGlySerGlnProMetSerProSerAspPheLeu 48
Db 481 ACGAAGGAAACAAACAGTCACGCGTCCGAGAGTCCCAATGTCACCTTCGATTTTGTG 540
OY 49 AsplysLeuMetGlyArgThrSerGlyThrAspAlaArgIleArgProAsnPhelysGly 68
Db 541 GACAAACTAATGGGAGACATCCGGATATGATCAAGAAATCAGGCCCACTTCAAAAGT 600
OY 69 ProProValAsnValThrCysAsnIlePheIleAsnSerPheSerValThrLysThr 88
Db 601 CCTCCAGTTAATGTCACATGCAACATTTTCAATAACACCTTGGCTCGCCGACAGAGC 660
OY 89 ThrIleAspTyrArgValAsnValPheLeuArgGlnGlnTyrPasnAspProArgLeuser 108
Db 661 ACTATGATTTACAGATAACATTTTCTCTGTCGAGAGTGGAATGATCCTCGCTTGA 720
OY 109 TyrArgGluTyrProAspAspSerLeuAspLeuAspProSerMetLeuAspSerIleTyr 128
Db 721 TACAGTGAATACCCGATGATTCATTAGACCTCCGCCATCATCTGTGACTCCATATAG 780
OY 129 LysProAspLeuPhePheAlaAsnGluTyrGlyAlaAsnPhelIleGlnValThrThrAsp 148
Db 781 AAACCTGCTGTTCTTGTCTAATGAGAAAGGGGCTTCTCCACGAGATCCGCCGAT 840
OY 149 AsnLysLeuLeuArgIlePheLysAsnGlnLysValLeuTyrSerIleArgLeuThrLeu 168
Db 841 AACCAAGCGCTAAGAAATTTTCAAAATGGAATGTTCTTTATTCATAAGGTGACATTA 900
OY 169 IleLeuSerCysLeuMetAspLeuLysAsnPhelProMetAspIleGlnThrCysThrMet 188
Db 901 ACACCTCTGTCCAATGATCTCAAGAAATTTCCCAATGATGTTCAAAACATGATATAG 960
OY 189 GlnLeuGluSerSerIleLeuCysSerProLeuProSerLeuSerLeuSerValGly 208
Db 961 CAACTCGAA-----AGCTTTGGG 978
OY 209 TyrThrMetLysAspLeuValPheGluTyrPheGluLysAspAlaProAlaValGlnValAla 228
Db 979 TACACGATGAATGATCTCATTTTGCATATGCGCAAGATGAAGACCA---CTACAACTGCCT 1035

```

```

OY 229 GluGlyLeuThrLeuProGlnPheIleLeuArgAspGluLysAspLeuGlyCysThr 248
Db 1036 GAAGCACTCACTTCCCTCAATTTCTGTGAAAGAAAGAAAGATTTGCCATCTGCACT 1095
OY 249 LysHisTyrAsnThrGlyLysPheThrCysIleGluValLysPheHisLeuGluArgGln 268
Db 1096 AAACACTCAATTAAGAGAAAGTTTACATGCTTAAAGATGACATTTCACTTGACGGCNA 1155
OY 269 MetGlyTyrTyrLeuIleGlnMetTyrIleProSerLeuLeuIleValIleLeuSerTrp 288
Db 1156 ATGGGCTACTACTGATGCCAATGATCATTTCCAGCCCTTGATGATTCATCTGCTCGG 1215
OY 289 ValSerPheTrpIleAsnMetAspAlaAlaProAlaArgValGlyLeuGlyIleThrThr 308
Db 1216 GTCTCATCTTGTGATTAACATGATGACAGCTCCGCTGGGAGGCTTGATTCACCAAGTCTCTAT 1275
OY 309 ValLeuThrMetThrThrGlnSerSerGlySerArgAlaSerLeuPhePolysValSerTyr 328
Db 1276 GTACTTACGATGACACCGACAGATTCGTGATCCGGGCTTCTTACCAAGGTCTCTAT 1335
OY 329 ValLysAlaIleAspIleTrpMetAlaValCysLeuLeuPheValPheAlaAlaLeuLeu 348
Db 1336 GTCAAGGCAATTTGACATTTGATGACAGTGTCTCTTTTGTGTTCTGACACTTCTG 1395
OY 349 GluTyrAlaAlaIleAsnPhelValSerArgGlnHisLysGluPheIleArgLeuArgArg 368
Db 1396 GAGTATGACGCCGGAATTTGTATCAAGGCCAACAAGAACTGAGGTTTGGCGGA 1455
OY 369 ArgGlnArgArgGln-----ArgLeuGlu----- 376
Db 1456 AAGAGGAAATTAACAGAAAGCTTTTGACACTGGAGAAAGTTTACCGTTTCTCAGACAGC 1515
OY 377 GluAspIleIleGlnGluSerArgPheThrPheArgGlyTyrGlyLeuGlyHisCysLeu 396
Db 1516 GATGATGAGTGAGAGGAGAGTCCGCTCAGCTTCACTGCTTATGAAATGAGGCCCTGCTT 1575
OY 397 GlnAlaArgAspGly-----GlyProMetGluGlySerGlyIleTyrSerPro 412
Db 1576 CAAGCAAGAGATGTGTGTTCCAAAGGTCCCAACCATGCTGTCCAGGTC----- 1626
OY 413 GlnProAlaProLeuLeuArgGluGlyGluThrThrArgLysLeuTyrValAsp 431
Db 1627 -----ATGCCAAGAGCGCCCATGAAATGAGGAAAGGCTTTCATGAC 1668

```

```

RESULT 12
HSGLYRA2 1715 bp mRNA linear PRI 28-MAY-1993
LOCUS H.sapiens alpha-1 strychnine binding subunit of inhibitory glycine
DEFINITION receptor mRNA.
ACCESSION X52009.1 GI:31850
VERSION X52009.1
KEYWORDS glycine receptor; inhibitory glycine receptor; strychnine binding.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1715)
AUTHORS Grenningloh,G., Schlieden,V., Schofield,P.R., Seeburg,P.H.,
Siddique,T., Mohandas,T.K., Becker,C.M. and Betz,H.
Alpha subunit variants of the human glycine receptor: primary
structures, functional expression and chromosomal localization of
the corresponding genes
EMBO J. 9 (3), 771-776 (1990)
JOURNAL MEDLINE 90183975
PUBMED 2155780
FEATURES
source
location/Qualifiers
1..1715
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/map="Xp21.2-p22.1"
/tissue.type="brain"
/dev_stage="fetal"

```

```
gene
297.1646
/ene="GLYRA2"
CDS
297.1646
/ene="GLYRA2"
/note="strychnine binding alpha-1 subunit"
/codon_start=1
/product="inhibitory glycine receptor"
/protein_id="CAA36258.1"
/db_xref="GI:31851"
/db_xref="SWISS-PROT:P23A15"
/translation="MSPNTLRKLYSGAIVFESLASKENAEARSATKPMSPDFDK
LMKRTSGDRLRPNKGPVNVSCNFINSEGSIAETWYRVNIFLRQWMDPRLA
YNEIPDSLDLPMSLMSIKPDLFANEAGAFHEITDKLRLSRNGVNLRI
TTLACPMYKPNPDVQICIMQLESFYMNDLIEEMOGAYSDGLTLPOLIK
EEKDLRCKHYNTGKTCIEARFHLEROMGYLIOMYIPSLIIVISMTSEPMINDA
APARVGLITTVLTMTTOSGSRSLPKSYVKAIDIMAVCLLPFSALEEYAVNF
VSROKHELRLFRKRRHKKHKKEDGEGRENFNSATGMCPACIQADGSIKVCANNSTNT
PPAPSKSPKMKLFIQRAKKIDKISRLGFPMAFLIFNFWIITKIVREDVHNO"
BASE COUNT 424 a 480 c 407 g 404 t
ORIGIN
Alignment Scores:
Pred. No.: 1.03e-150 Length: 1715
Score: 1560.50 Matches: 307
Percent Similarity: 82.54% Conservative: 38
Best Local Similarity: 73.44% Mismatches: 50
Query Match: 70.18% Indels: 23
DB: Gaps: 7
US-10-075-846-4 (1-431) x HSLYRA2 (1-1715)
QY 17 LeuProGlyGlnValLeuAlaLeuAlaLeuGlnValLysSerGlyThr 36
Db 327 CTTTCGGAGCCATTGATCTTCAGCCTTGCTCTTAAGAGGTGAAGCTGCC 386
QY 37 LysGlySerGlnProMetSerProSerAspPheLeuAspLysLeuMetGlyArgThrSer 56
Db 387 TCCGCAACCAAGCCTATGTCACCTCGATTCCTGATTAAGCTAAATGCGAGCAACCTCC 446
QY 57 GlnTyraAspAlaArgIleArgProAsnPheLysGlyProProValAsnValThrCysAsn 76
Db 447 GGAATGATGCCAGATGACGAGCAATTTAAAGTCCCGCAGGAAAGCGATGCGTGAAC 506
QY 77 IlePheIleAsnSerPheSerValThrLysThrMetAspTyrArgValAsnVal 96
Db 507 ATTTGATCAACAGCTTGCTTCATTCGACGAGCAACCATGAGCTTAAGGTCAACATC 566
QY 97 PheLeuArgGlnGlnIntrpAsnAspProAlaGLeuSerTyrArgLutyrProAspSer 116
Db 567 TTCCTGGCGAGCATGGAAGCAACCCCGCTGCTATATGAATACCTGACGACTCT 626
QY 117 LeuAspLeuAspProSerMetLeuAspSerIleTyrLysProAspLeuPhePheAlaAsn 136
Db 627 CTGGACCTGGACCATTCATGCTGAGACTCATCTGGAACCTGACCTGTTCTTGGCCAAC 686
QY 137 GlnLysGlnAlaAsnPheHisGlnValThrThrAspAsnLysLeuLeuArgIlePheLys 156
Db 687 GAGAGGGGGCCCATTCATGAGATACACACAGCAAAATTCGTAAGATCTCCCG 746
QY 157 AsnGlnAsnValLeuTyrSerIleArgLeuThrLeuIleLeuSerCysLeuMetAspLeu 176
Db 747 AATGGCAATGTCTCTACACATCAGAAATCACCTGACCTGGCTGCCCATGAGACTTG 806
QY 177 LysAsnPheProMetAspIleGlnThrCysThrMetClnLeuGlnSerSerIleLeu 196
Db 807 AAGAAATTTCCCATGATGTCAGACATGTATCATGCAACTGGAA----- 851
QY 197 CysSerProLeuProSerLeuSerLeuSerValGlnTyrThrMetLysAspLeuValPhe 216
Db 852 -----ACCTTGGAATATACGATGATGACATGACCTCATCTTT 884
QY 217 GlnTyrLeuGlnAspAlaProAlaValGlnValAlaGlnLysLeuThrLeuProGlnPhe 236
||||| ||| |||||||||||||||||||||||||||||||||||||||
```

```
Db 885 GAGTGGCAGAGAACAGGGA---GCCGTGACAGTATGACAGATGAGACTACTGCCCCAGATT 941
QY 237 IleLeuArgAspGlnLysAspLeuGlnCysCysThrLysHisThrAsnThrGlyLysPhe 256
Db 942 ATCTTGAAAGAAAGAAAGAGCTTGAGATCTGACACCAAGACACTACACAGGTAATTC 1001
QY 257 ThrCysIleGlnValLysPheHisLeuGlnArgGlnMetGlyTyrTyrLeuIleGlnMet 276
Db 1002 ACTGCAATGAGCCCGGTTCCACCTGAGAGGAGGACATGGGTATGACATGATTCAGATG 1061
QY 277 TyrIleProSerLeuLeuIleValIleLeuSerTyrValSerPheThrPheAsnMetAsp 296
Db 1062 TATATTCGCCAGCTGCTCATGATTCATCCCTCATGATGATCTCTTGGATCAACAGGAT 1121
QY 297 AlaAlaProAlaArgValGlyLeuGlyIleThrThrValLeuThrMetThrPheGlnSer 316
Db 1122 GCTGCACCTGCTGTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 1181
QY 317 SerGlySerArgLysSerLeuProLysValSerTyrValLysAlaIleAspIleThrMet 336
Db 1182 TCCGGCTTCGACATCTCTGCCCAAGGTGCTCTATGTGAAGCAATTCATTTGGATG 1241
QY 337 AlaValCysLeuLeuPheValPheAlaAlaLeuLeuGlnTyrAlaAlaIleAsnPheVal 356
Db 1242 GCAGTTGCTGCTGCTTGTGTTCTCAGCCCTATTGAATATGCTGCCCTTAATCTTTGTC 1301
QY 357 SerArgGlnHisLysGlnPheIleArgLeuArgArgGlnArgGlnArgGlnArgLeuGln 376
Db 1302 TCTCGCAACATTAAGAGGCTGCTCCGATTCAGAGAGGAGGAGACATCAACAG----- 1355
QY 377 GlnAspIleIleGlnLysSerArgPheTyrPheArgGlyTyrGlyLeuGly---HisCys 395
Db 1356 GAGAGTGAACCTGAGAGAGCGCTTAATCTCTGCGCATATGAGATGGGCCAGCCTGT 1415
QY 396 LeuGlnAlaArgAspGlyGlyPrometGlyGly-----SerGlyIleTyrSerPro 412
Db 1416 CTACAGGCCCAAGATGATGATCTGACCAAGCGCCCAACAAACCAACCAACCAACCCC 1475
QY 413 GlnProProAlaProLeuLeuArgGlnGlyGlnThrThrArgLysLeuTyrVal 430
Db 1476 ---CCTCTGACCA---TCTAAGTCCCGCAGAGAGATGCAAAACTCTTCATC 1523
RESULT 13
AF268375 1744 bp mRNA linear MAM 22-MAY-2001
LOCUS Bos taurus glycine receptor alpha 1 subunit (G1ra1) mRNA, complete
DEFINITION Bos, alternatively spliced.
ACCESSION AF268375
VERSION AF268375.1 GI:10180958
KEYWORDS
SOURCE
ORGANISM Bos taurus.
Bos taurus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 1744)
Pierce, K.D., Handford, C.A., Morris, R., Vafa, B., Dennis, J.A.,
Healy, P.J., and Schofield, P.R.
A nonsense mutation in the alpha1 subunit of the inhibitory glycine
receptor associated with bovine myoclonus
Mol. Cell. Neurosci. 17 (2), 354-363 (2001)
JOURNAL MEDLINE 21109390
PUBMED 11178872
REFERENCE 2 (bases 1 to 1744)
Pierce, K.D., Handford, C.A., Morris, R., Vafa, B., Dennis, J.A.,
Healy, P.J., and Schofield, P.R.
Direct Submission
Submitted (17-MAY-2000) Neurobiology Program, Garvan Institute of
Medical Research, 384 Victoria Street, Sydney, NSW 2010, Australia
Location/Qualifiers
1. 1744
/db_xref="taxon:9913"
FEATURES
SOURCE
```



```
gene
1. 1744
/gene="Glr1"
229. 1602
/gene="Glr1"
/note="Glr1ins; alternatively spliced"
/codon_start=1
/product="glycine receptor alpha 1 subunit"
/protein_id="AAG14346.1"
/db_xref="GI:10180959"
/translation="MSEPTLRIVIMENTIVFSLAASKAEASAKSPSPDLK
LMGRSGYDARIRPFKGPYVNSCNTINFSGLAETTMRYRNIPIRQOMDBRLA
INERPDSDLDIPSMDSIMKPDLEFANEKGAHFHEITDKLISNGVNLSTIRI
TLTLACPMDLNLFPMDOVOTCMOLESEGYTNDLIFEMOEGAAVOVADGLPQILK
EKEDIRYCTKHNTGKFTFCIEARPHLEKMGVLYIOMYIPSLIYIISWISFWINDA
APARVGLGITVLTQTOSGSRASLPVSYVKAIDIMAVCLLEFSALEAVANF
VSRHKKELRRRRKRRHKSPLNLFODEAGEGENFSEAYGMPACIOARDGISVKG
ANNSTNTPPAPSPKSPKEMKLFTRKAKIDKISRIGFPAFLIFNMFTIYKIVR
REDVNO"
BASE COUNT      396 a      508 c      451 g      389 t
ORIGIN
Alignment Scores:
Pred. No.:      1,19e-150      Length:      1744
Score:          1580.00      Matches:      309
Percent Similarity: 80.89%      Conservative: 38
Best Local Similarity: 72.03%      Mismatches: 52
Query Match:    70.16%      Indels:      30
DB:            4      Gaps:      8
US-10-075-846-4 (1-431) x AF268375 (1-1744)
OY 12 LeuleuLeuTPrThLeuProGlyValLeuLeuArgValAlaLeuAlaLysGlu 31
DB 253 CTCTACCTTTGGAG-----ACCATGTATCTTCTCAGTCTGTCTGCTTCCAAAG 303
OY 32 ValLysSerGlyThLysGlySerLInPrometSerProSerAspPhaLeuAspLysLeu 51
DB 304 GCTGAAGCTGCTCGCTCTGCTTCCAAAGCCAGTCACGCTCGATTTCTCGATTAACATC 363
OY 52 MetGlyArgThrSerGlyTyrAspAlaArgLeuArgProAspPhaLysGlyProProVal 71
DB 364 ATGGGAGAGACTCTCGATACGACCCAGCATGAGCCCATTTCAAGAGTCCCAATG 423
OY 72 AsnValThrCysAsnIlePheIleAsnSerPheSerValThrLysThrMetAsp 91
DB 424 AATGTACACTGCACATTTTCATCAACAGCTTCGCTTCATTCATGAGAACACTATGAC 483
OY 92 TyrArgValAsnValPheLeuArgGlnGlnTrpAsnAspProArgLeuSerTyrArgGlu 111
DB 484 TACAGGTCACATCTTCTCGGCGACAGTGAATGACCCCGGCTGCTACATGAA 543
OY 112 TyrProAspSerSerLeuAspLysAspProSerMetLeuAspSerIleTrpLysProAsp 131
DB 544 TACCCCGATGACTCGCTGAGCTTGAGCCATCCATGTCATCTGGAACACTGAC 603
OY 132 LeuPhePheAlaAsnGlyLysGlyAlaAsnPheHisGluValThrThAspAsnLysLeu 151
DB 604 CTGTTCTTTGCCAAAGAGAGGGGGCCCATCTTCATGAGATCACACAGACAACTAATG 663
OY 152 LeuArgIlePheLysAsnGlyAsnValLeuTyrSerIleArgLeuThrLeuIleLeuSer 171
DB 664 CTGAGGATCTCCCGAAGCGGAATCTCTCTACACCATTTAGATACCCCTGACATTTGGCC 723
OY 172 CysLeuMetAspLeuLysAsnPheProMetAspIleGlnThrCysThrMetGlnLeuGlu 191
DB 724 TGCCCATGAGACTTAAGAAATTTCCCATGATGATCCAGAGTGCATTAATGAACTGAA 783
OY 192 SerSerSerIleLeuCysSerProLeuProSerLeuSerValGlyTyrThrMet 211
DB 784 -----ACCTTGGATACACCATG 801
OY 212 LysAspLeuValPheGluTrpLeuGluAspAlaProAlaValGlnValGlnGlyLeu 231
|||||
```

```
DB 802 AACGACCTACTTTGAGTGGCAGGAGAGGG---GCTGTGACAGTGGCAGATGACTA 858
OY 232 ThrLeuProGlnPheIleLeuArgAspGluLysAspLeuGlyCysCysThrLysHisTyr 251
DB 859 ACTTACCCCAAGTTTATCTTCAAGAGAGAGAGACTTGAATACTGCACCAACATTTAC 918
OY 252 AsnThrGlyLysPheThrCysIleGluValLysPheHisLeuGluArgGlnMetGlyTyr 271
DB 919 AACACAGGCAAAATTCACCTGCATCGAGGCCGGGTTCACCTCGAGCGACAGATGGCTAC 978
OY 272 TyrIleuIleGlnMetTyrIleProSerIleLeuIleValIleLeuSerTPrValSerPhe 291
DB 979 TACCGATTCAGATGATCATCTCCAGCCTGCTCATCTTCATCTCTCGTGGATCTCTTC 1038
OY 292 TrpIleAsnMetAspAlaAlaProAlaArgValGlyLeuGlyIleThrValLeuThr 311
DB 1039 TGGATCAACATGATGCTGCTCCGCCCGCTGGGCTGGGCATCACCATGCTCTCAC 1098
OY 312 MetThrThrGlnSerSerGlySerArgAlaSerLeuProLysValSerTyrValLysAla 331
DB 1099 ATGACACACCCAGAGCTCGGCTCGAGCATCCCTGCCCAAGTATGATGAGAGGCC 1158
OY 332 IleAspIleTPrMetAlaValCysLeuLeuPheValPheAlaAlaLeuLeuGlyTyrAla 351
DB 1159 ATCGACATTTGGATGGCCGCTGCTGCTCTTGTGTTCTCAGCCCTGCTTACAGCC 1218
OY 352 AlaIleAsnPheValSerArgGlnHisLysGlnPheIleArgLeuArgArgArgGlnArg 371
DB 1219 GCCGCAACTTTGTGTCTCGGCAGACAAAGCTGCTCCATTCAGAGAGAGAGAGAGA 1278
OY 372 ArgGlnArg-----LeuGluGluAspIleIleGlnIleSerArgPhe 385
DB 1279 CATCACAGAGACCCCATGTAATCTGTTCCAGAGAGATGAGCGTGGAGAGCGCCCTTC 1338
OY 386 TyrPheArgGlyTyrGlyLeuGly---HisCysLeuGlnAlaArgAspGlyLysPromet 404
DB 1339 AACTCTCGCCCTAGCGATGGGCGCCAGCTGCTTCCAGGCAAGAGGATCTGCTGTC 1398
OY 405 GluGly-----SerGlyIleTyrSerProGlnProAlaProLeuLeuArgGlu 421
DB 1399 AAGGCGGCCAACAACAGACAGACCAACCT---CCTCGGAGACCA---TCCAAAGTCC 1452
OY 422 GlyIleThrThrArgLysLeuTyrVal 430
DB 1453 CCGGAGAGATGCGAATAACTCTTCATC 1479
RESULT 14
AF362764 1392 bp mRNA linear ROD 09-AUG-2001
LOCUS Mus musculus glycine receptor alpha 3 subunit (Glr3) mRNA,
DEFINITION complete cds.
ACCESSION AF362764
VERSION AF362764.1 GI:14010348
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1392)
AUTHORS Noegel,S., Becker,C. and Becker,K.
TITLE Different glycine receptor isoforms are expressed in murine
cerebellum
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1392)
AUTHORS Noegel,S., Becker,C. and Becker,K.
TITLE Direct Submission
JOURNAL Submitted (21-MAR-2001) Biochemistry, University of Erlangen,
Fahrrstrasse 17, Erlangen 91054, Germany
FEATURES
source
1. 1392
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
```

```

/chromosome="8"
/map="26 cm"
/tissue_type="spinal cord"
1. .1392
/gene="Glr3a"
1. .1392
/gene="Glr3a"
/note="Ligand-gated ion channel"
/product="glycine receptor alpha 3 subunit"
/protein_id="AAK51962.1"
/db_xref="GI:14010349"
/translation="MAHVRHPRLLVSGFYWEAALLSLVATKETNSARSRAPMSGS
DFLDKLMGRISGYDARIKPNFKGPVNVNTNIFINSFGSIATETMDRVNIFLRQKN
DPRLAYSPDSDLDPSMDSIMKPDLEFANEKGFHEVTDNKLIFKNGNVLV
SIRLTLLSCPMDLKPNPDVOTCINQJESFGYTMNDLIFEMODEAPVOYAEGLTLPQ
FLKEEKDLRYCTKRYNTGKTECFEVRFHLEROMGYLIOMYIPLSILIVLSWSPFI
NMADAPARVALGITTDLDTTQSSGSNASLPKSYVAIDIMAVCLLFYSALLEIA
AVNFVSRQHKELFRKRKRKNTKTEAFLEFVRSDDVDEVRSEFETAVGMPCLQ
AKGVVPRGNHAAVQVMPKSPDEMRKVFIDRAKKIDTISRACEPLAFLINIFVWVIY
KILRHEDIHQOO"
BASE COUNT      388 a      310 c      313 g      381 t
ORIGIN
Alignment Scores:
Pred. No.:      1,25e-150      Length:      1392
Score:          1578.50      Matches:      317
Percent Similarity: 79.19%      Conservative: 33
Best Local Similarity: 71.72%      Mismatches: 50
Query Match:      70.09%      Indels:      42
DB:              10      Gaps:      9
US-10-075-846-4 (1-431) x AF362764 (1-1392)
OY      8 ThrLeuSerPheLeuLeuLeuTrp---ThrLeuProGlyGln-ValLeuLeuArgValAl 26
    ||||| ||| ||| |||||
Db      15 ACACCTTCGGACACTACTTCTTGATTTTACTTCTGGAAAGCCGCACTGTACTGACTTT 74
OY      26 aLeuAlaLysGlnGluValLysSerGlyThrLysGlySerGlnProMetSerProSerAs 46
    ::||| ::||| ||| ||| ||||| |||||
Db      75 GGTCGCCACAAGAAACAACAGTGGCGCATCGCAAGTGCCTCCATGTCACCTTCTGA 134
OY      46 PheLeuAspLysLeuMetGlyArgThrSerGlyTyrAspAlaArgIleArgProAsnPh 66
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      135 TTTTCTGCAAACTAAATGGGAGACATCGGGGTATGATGCAAGAATCAGACCACACTT 194
OY      66 eLysGlyProProValAsnValThrCysAsnIlePheIleAsnSerPheSerValTh 86
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      195 CAAGGTCCTCCAGTAAATGTCACATGCAACATATTCATTAACACAGCTTTGGCTCCATTGC 254
OY      86 rLysThrThrMetAspTyrArgValAsnValPheLeuArgGlnGlnIleTrpAsnAspProAr 106
    ::||| ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      255 AGAGACGACATATGATATACAGAGTAACATTTTCTTCGTCAGAAAGTGAATGATCTCG 314
OY      106 GLeuSerTyrArgGlyTyrProAspAspSerLeuAspLeuAspProSerMetLeuAspSe 126
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      315 TCTTGCACTAC---AGTATATCTGACGATTCATTAGACCTTGACCCGCTATGTTGATTC 371
OY      126 rIleTrpLysProAspLeuPhePheIleAsnGlnLysGlyValAsnPhenHisGluValTh 146
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      372 CATTGGAAAGCTGACTTGTCTTCTTAATGAAGAAGGGGCTTAACCTTCCAGGAAGTCAC 431
OY      146 rThrAspAsnLysLeuLeuValrGlePheLysAsnGlyAsnValLeuTyrSerIleArgGle 166
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      432 CACAGATTAACAAGCTGCTAAGATTTTCAAAAATGGAAGTCTCTTATTATCAATAAGTTC 491
OY      166 uThrLeuIleLeuSerCysLeuMetAspLeuLysAsnPhenProMetAspIleGlnThrCy 186
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      492 GACATTTACACTCTCCCTCCAAATGATCTCAAAATAATTTCCCATGATGATACAAACATG 551
OY      186 sThrMetGlnLeuGlnSerSerIleLeuCysSerProLeuProSerLeuSerLeuSe 206
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      552 CATATATCAACTTGA-----AG 569

```

```

OY      206 rValGlyTyrThrMetLysAspLeuValPheGlyTrpLeuGluAspAlaProAlaValAl 226
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      570 CTTTGGGTACACATGATGATGTCATTCTTGAATGACAGATGAAGACCCCA---GTACA 626
OY      226 nValAlaGlyLeuThrLeuProGlnPheIleLeuArgAspGluLysAspLeuGlyCy 246
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      627 AGTGGCGAAGGACTCTCTTGGCTCAATTCCTGTGAAGAAAGAAAGGACTTGGCGTA 666
OY      246 sCysThrLysHisThrAsnThrGlyLysPheThrCysIleGluValLysPheHisLeuGl 266
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      687 CTGCACATAACACTACATATACAGGAAGTTTACTGATGATGAAAGTGCGATTCCTTGA 746
OY      266 uArgGlnMetGlyTyrTyrLeuIleGlnMetCysIleArgProSerLeuLeuValIleLe 286
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      747 GCGTCATATGGGCAATTAATCAATGACAGATGATATCCAGCCAGCTCTGATTCGTTATTC 806
OY      286 uSerTrpValSerPheThrPheLeuAsnMetAspAlaAlaProAlaArgValIleGlyI 306
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      807 ATCTGGCTCTCATCTTGATTAACATGATGACAGCCCGCTAGACTTGGGCGAT 866
OY      306 eThrThrValLeuThrMetThrThrGlnSerSerGlySerArgAlaSerLeuProLysVa 326
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      867 CACCACGATCTTACTATGACACACAGACTTCTGATCCCGGCGCTCTTACCAAGAT 926
OY      326 lSerTyrValLysAlaIleAspIleTrpMetAlaValCysLeuLeuPheValPheAlaAl 346
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      927 GTGATACGTCACAGGCTATGACATTTGATGAGCGGTGCTGCTCTTTGTGTTCTACGC 986
OY      346 aLeuLeuGlyTyrAlaAlaIleAsnPheValSerArgGlnHisLysGluPheIleArgLe 366
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      987 ACTTCTGAGTATGACACCGCTGATTTTGTCTCAAGGCAACACAGAGAGCTTTGAGCTT 1046
OY      366 uArgArgArgGlnArgArgGln-----ArgLeuGlu----- 376
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1047 TCGCGCAAGAAAGAAATGAAGACGGAAGCTTTTGCATGAGAGATTTTACCCTTTCTC 1106
OY      377 -----GluAspIleIleGlnIleuSerArgPheTyrPheArgGlyTyrGlyLeuGlyH 394
    ::||| ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1107 AGACACGAGATGATGAGTGAAGGAGACTGATTCAGCTTCACTGACCTCATGGAATGGGGCC 1166
OY      394 sCysLeuGlnAlaArgAspGly-----GlyProMetGlnGlySerGlyIleTyr 410
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1167 CTGCTCTCAAGCCCAAGATGATGTTGTTCCAAAAGGCCCAACACACCTCTCCAGCTC-- 1224
OY      410 rSerProGlnProProAlaProLeuLeuArgGlyGlyGlnThrThrArgLysLeuTyrVa 430
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1225 -----ATGCAAGAGAGCCCT-----GATGAATAGGAAGGCTTTCAT 1262
OY      430 lAsp 431
    |||||
Db      1263 CGAC 1266

```

RESULT 15  
 DRES812  
 LOCUS DRES812 1628 bp mRNA linear VRT 01-AUG-1998  
 DEFINITION Danio rerio mRNA for glycine receptor (alpha1 subunit).  
 ACCESSION AJ005812  
 VERSION AJ005812.1 GI:3378595  
 KEYWORDS alpha1 subunit; glycine receptor.  
 SOURCE zebrafish.  
 ORGANISM Danio rerio  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi;  
 Cypriniformes; Cyprinidae; Danio.  
 REFERENCE 1 (bases 1 to 1628)  
 AUTHORS David Watline,B.  
 TITLE Direct Submission  
 JOURNAL Submitted (27-APR-1998) David Watline B., Biotechnologies, Inserm  
 U261 - Institut Pasteur, 25, Rue du Dr. Roux, 75724 Paris cedex15,  
 FRANCE  
 2 (bases 1 to 1628)  
 REFERENCE David Watline,B., Goblet,C., de Saint Jan,D., Fucille,S.,  
 AUTHORS



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: June 30, 2003, 21:13:09 ; Search time 345.186 Seconds  
(without alignments)  
2811.853 Million cell updates/sec

Title: US-10-075-846-4  
Perfect score: 2252  
Sequence: 1 MTLVPATLSFLMTLMTLPCQ.....POPPAPALRGCTTRKLYVD 431

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues  
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL-frame+ p2n model -DEV-xip  
-O/cgrr2\_1/USPTO.spool/US10075646/runtat\_25062003\_163647\_5141/app-query.fasta.1.782  
-DB-N.Geneseq.101002 -QFMT-fastlap -SUFFIX-p2n.rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS-bits -START=1 -END=1 -MATRIX-blosum62 -TRANS-human40.cd1  
-LIST=45 -DOCALLIG=200 -THR\_SCORE-pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE-LOCAL -OUTFMT-pto -NOM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER-US10075646.ecgn.1.1.511.arnat.25062003.163647.5141 -NCPu=6 -ICPp=3  
-NO\_MMAP -LAREQUIDRY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGCLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
1: N.Geneseq.101002:\*  
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*  
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*  
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*  
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*  
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*  
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*  
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*  
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*  
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*  
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*  
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*  
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*  
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*  
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*  
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*  
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*  
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*  
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*  
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*  
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*  
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*  
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*  
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*  
25: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2163	96.0	1874	24 AAD33667	Human TRICH-22 CDN
2	2157	95.8	1600	21 AAC61678	CDNA sequence enco
3	984.5	43.7	657	13 AAD25722	GABA-A receptor be
4	871	38.7	2404	24 AB199252	Mouse ischaemic co
5	761	33.8	39796	21 AAC61681	Nucleotide sequenc
6	754.5	33.5	2400	22 AAD21397	R. sanguineus glut
7	750	33.3	4621	24 AAD26939	Heliothis virescen
8	749	33.3	1609	24 AAD26937	Heliothis sp. HEC3
9	742.5	33.0	1197	22 AAD21378	R. sanguineus LGIC
10	742.5	33.0	1197	22 AAD22082	R. sanguineus Gluc
11	742.5	33.0	2138	22 AAD21395	R. sanguineus glut
12	742.5	33.0	2289	22 AAD21396	R. sanguineus glut
13	737.5	32.7	1614	24 AAD22072	Dermacentor variab
14	737	32.7	1614	24 AAD22070	Dermacentor variab
15	732.5	32.5	1614	24 AAD22071	Dermacentor variab
16	731.5	32.5	1368	20 AAX24372	Cat flea glutamate
17	731.5	32.5	3958	17 AAT4356	Glutamate-gated ch
18	729	32.4	2340	23 AB116609	Drosophila melanog
19	728.5	32.3	1640	24 AAD26938	Heliothis sp. HECM
20	728.5	32.3	1844	20 AAV64372	GABA-gated chlorid
21	726	32.2	2066	21 AA244048	D. simulans GABA r
22	725.5	32.2	1766	22 AAD04276	Short form of S. a
23	725.5	32.2	1766	22 AAD04277	Short form of S. a
24	724.5	32.2	1519	20 AAV64373	GABA-gated chlorid
25	724	32.1	1657	19 AAV55007	GABA-gated chlorid
26	724	32.1	1657	19 AAV55008	GABA-gated chlorid
27	720.5	32.0	3619	23 AB113035	Drosophila melanog
28	720	32.0	2066	21 AA244045	D. melanogaster po
29	720	32.0	2066	21 AA244046	D. melanogaster po
30	719	31.9	3442	22 AAD21373	Dermacentor variab
31	719	31.9	3598	22 AAD21372	Dermacentor variab
32	718	31.9	2066	21 AA244047	D. melanogaster GA
33	717.5	31.9	1823	19 AAX16065	DNA sequence of lu
34	717	31.8	1491	17 AAO99330	GABA receptor subu
35	717	31.8	1491	20 AAV84665	Insect GABA recept
36	715	31.7	1498	23 AB118879	Drosophila melanog
37	712	31.6	1866	14 AAO33131	GABA-A receptor be
38	712	31.6	1866	15 AAO69143	Human GABA recept
39	698.5	31.0	1640	22 AA160040	Human polynucleoti
40	698.5	31.0	1693	22 AA158254	Human polynucleoti
41	687.5	31.0	1594	24 AB199308	Mouse ischaemic co
42	694.5	30.8	1866	24 ABL67062	Thyroid cancer rel
43	690	30.6	1879	22 AAD04274	Long form of S. am
44	690	30.6	1879	22 AAD04275	Long form of S. am
45	687.5	30.5	1970	21 AA244049	D. melanogaster po

## ALIGNMENTS

RESULT 1  
AAD33667  
AAD33667 standard; CDNA: 1874 bp.  
01-JUL-2002 (first entry)

Human: transporter and ion channel; TRICH-22; transport disorder; angina;  
amyotrophic lateral sclerosis; cystic fibrosis; neuromuscular disorder;  
cardiac disorder; polymyositis; diabetes; neurological disorder; cancer;  
depression; schizophrenia; anaemia; Wilson's disease; Cushing's disease;  
cell proliferated disorder; infertility; arteriosclerosis; gene therapy;  
Alzheimer's disease; Parkinson's disease; Huntington's disease; allergy;  
myasthenia gravis; multiple sclerosis; metabolic disorder; hypertension;  
acquired immune deficiency syndrome; immunological disorder; scleroderma;



Db	964	GAGAAAGATCATAGGCTGTGGACCAACGACTACAACAACAGGAAATTCACCTGCATGCAG	1023
OY	261	Vallysphenhstleugluaraglmeglytyrtyrleuileglnmettyrileproser	280
Db	1024	GTTAAAGTTTCAACCGAAGCGCAGATGGCGTACATATCTGATTCAGATGTACATCCCCAGC	1083
OY	281	Leuenuilevalileleuserrtrivalsepherprilleasnmelaspalaalproala	300
Db	1084	CTACTCATCGTCATCCGTCTGCTGGGTCTCCTTCTTGAGTACACATGATGCTGCTGCC	1143
OY	301	ArgvalglyleuglylierthrthrvallleuthrmethrthrnglnserSerglySerarg	320
Db	1144	CGTGTGGCGCTGGGCATCACACCGCTGCTCACATGACACACCCAGAGCTGTGCTCCGG	1203
OY	321	AlaserleuprolysvalsertryvallysalileaspilietrypmelalavalCysleu	340
Db	1204	GCCTCTTGGCTTAAGTGCTCCTACGTGAAGCAGATCGACATCTGGATGGCTGTGTCGTG	1263
OY	341	LeuphetvalphealalealeuleuglutrvalalaleasnphievalSerArgGlnhis	360
Db	1264	CTCTTTGTGTGGCTGGCTGCTGTGAGTATGCTCCATAAATTTTGTTCCTCGCAGCAT	1323
OY	361	LysgluetherleatrgleuArgArgGlnargrglnargrgleugluaspillele	380
Db	1324	AAAGAAATTCATTACGACTTCGAAAGAGCAGAGCGGCCACACCTTGGAGAAATATTCATC	1383
OY	381	GlnGlnuSerarqphetryrpheargGlytyrGlyleuglnhisCysleuGlnalAargasp	400
Db	1384	CAAGAAAGCTGTTCTCATTTCCGCGCATGTGGCTTGGGCCACTGCTGCACAGCAAGAT	1443
OY	401	GlyglytrpmetgluglyserglylietyserrproGlnProproAlaproleuLeuarg	420
Db	1444	GGAGGTCCAAAGAGGTCTGGCAATTAATAGTCCCAACCTCGAGGCCCTTCTTAAGG	1503
OY	421	GluGlygluThrthrArqlysLeuTyValasp	431
Db	1504	GAAGGAGAAACACGCGGAATCTACTGTGAC	1536
RESULT 2			
AAC61678			
ID	AAC61678	standard; DNA; 1600 BP.	
XX			
AC	AAC61678:		
XX			
DT	19-FEB-2001	(first entry)	
XX			
DE	cdna sequence encoding a human ataxia protein.		
XX			
KM	Human; ataxia; gene therapy; ss.		
XX			
OS	Homo saplens.		
XX			
FT	Key CDS	location/Qualifiers	
FT		1..1254	
FT		/*tag= a	
FT		/product= "ataxia protein"	
XX			
PN	MO200058461-A1.		
XX			
PD	05-OCT-2000.		
XX			
PE	23-MAR-2000; 2000MO-EP02600.		
XX			
PR	26-MAR-1999; 99EP-0106343.		
XX			
PA	(RAPP/) RAPPOLD-HOERBRAND G.		
XX			
PI	Rappold-Hoerbrand G;		
XX			
DR	WPI; 2000-656166/63.		
XX			
DR	P-PSDB; AAB19336.		
XX			
FT	Novel nucleic acid sequence encoding human ataxia protein for screening		

PT	compounds useful for treating disorders relating to mutations in ataxia gene -					
XX	Claim 2; Page 18-20; 47pp; English.					
PS						
XX	The present sequence encodes a human ataxia protein. The ataxia					
CC	protein and polynucleotides are useful for diagnosing and treating					
CC	disorders related to ataxia. Ataxia gene sequences are useful in					
CC	gene therapy, and as diagnostic tools or reagents for identifying and					
CC	characterizing genetic defect involved in the disorders and diseases					
CC	related to ataxia.					
XX						
SQ	Sequence 1600 BP; 378 A; 440 C; 379 G; 403 T; 0 other:					
<hr/>						
Alignment Scores:						
Pred. No.:	4,06e-240	Length:	1600			
Score:	2157.00	Matches:	414			
Percent Similarity:	96.29%	Conservative:	1			
Best Local Similarity:	96.06%	Mismatches:	2			
Query Match:	95.78%	Indels:	14			
DB:	21	Gaps:	1			
<hr/>						
US-10-075-846-4 (1-431) x AAC61678 (1-1600)						
OY	1 MetThrThrLeuValProAlaThrLeuSerPheIleuLeuLeuTrpHisLeuProGlyGln	20				
Db	1 ATACCAACTCTGTTCCTGCACAACCCTCCTTCCCTTCTCGAGCCCTGCCAGGGCCAG	60				
OY	21 ValIeuLeuArgValAlaLeuAlaLysGluGluValLysSerGlyThrLysGlySerGln	40				
Db	61 GTCTCTCTCAGGGTGCCCTTTGGCAAAGAGAAGTCAAAATCTGAACCAAGGGTCCCAG	120				
OY	41 PromEterProSerAspPheLeuAspLysLeuMetGlyArgThrSerGlyTyraAspAla	60				
Db	121 CCCATGCCCCCTCATATTCTTAGACAACATTAGGGGGCGAACATCTGATATGATGCC	180				
OY	61 ArgIIeaIrProAsnPhelSgLyProProValasnaValmrcyAsnIlePheIleasn	80				
Db	181 AGGATTTCGGCCCAATTTTAAGGCCACCCTCGAAGCGACCTGCACATCTTCATCAC	240				
OY	81 SerPheSerSerValmThrLysThrThrmetLaspyrArgValasnaValPheLeuArgGln	100				
Db	241 AGTTTCAGCTCCATTCACCAAGACCAACAAAGGCTACCGGGTGAATGCTTCTTCGGCAA	300				
OY	101 GlnTrpAsnAspProArgLeuSerTyraGrgIuTyrrProAspAspSerLeuAspLeuasp	120				
Db	301 CAATGGAAATGACCCACAGCCTCTCTACCAAGAAATTCCTGATGACTCTCTGGACTCGAT	360				
OY	121 ProSerMetLeuAspSerIleTrpLysProAspLeuPhePheAlaasnGluLysGlyAla	140				
Db	361 CCTTCATGCTGGAGCTCTATCTGGAAGCCACAGACCTCTTCTTGCTAATGAGAAGGGGCC	420				
OY	141 AsnPheHISgluValIthrThrAspAsnLysLeuLeuArgIlePheLysAsnGlyAsnval	160				
Db	421 AACTTCCATGAGGTGACCAAGCAACAAGATTACGCGCATCTTCACAAAGGGAAATGTG	480				
OY	161 LeuTySerIleArgLeuThrIleuIleuSerCySLeumetAspLeuLysAsnPhPro	180				
Db	481 CTGTACAGCATCAGGCTGACCTCATTTTGTCTCTCGATGAGACTTCAACAACACTTCCC	540				
OY	181 MetAspIleGlnThrCysThrmetGlnLeuGlnUserSerSerIleLeuCysSerProLeu	200				
Db	541 ATGAGCATCCAGACGCGACAGATGCGAGCTTGA-----	572				
OY	201 ProSerLeuSerLeuSerValGlyTyrrThrMetLysAspLeuValPheGlnTrpLeuGlu	220				
Db	574 -----AGTTTGGCTACACCAAGAAAGACCTCGTTTTGAGTGGCTGAA	618				
OY	221 AspAlaProLaValaJinValAlaGluGluLeuThrLaspProGlnPheIleLeuArgAsp	240				
Db	619 GATGCTCTCGCTCTCCAAAGTGGCTGAGGGGCTGACTTGCCTCCACAGTTTATCTTCGGGAT	678				
OY	241 GluLysAspLeuGlucyScyThrLysHisTyraSnThrGlyLysPheThrCysIleGlu	260				



```
Db 679 GAGAGAGATCTAGGCTGTGTACCAACACACACAGGAAATTCACCTGCATCGAG 738
Oy 261 ValysPheHisLeuArgGlnMetGlyTyrTyrLeuIleGlnMetTyrIleProSer 280
Db 739 GTAAAGTTTCCCTGGAAAGCGACATGGCTACTATCTGATTCAGATGATACATCCAGC 798
Oy 281 LeuLeuIleValIleLeuSerTyrValSerPheThrPheLeuMetAspAlaIleProAla 300
Db 799 CTACTCATCGCATCGTCTGCTGGGTCTCTTCTGATCAGCATGATGCTGCCCTGCC 858
Oy 301 ArgValGlyLeuGlyIleThrValLeuThrMetThrThrGlnSerSerGlySerArg 320
Db 859 CGTGGGCGCTGGGACATCACCGCTGCTCAGCATGACACACAGAGCTGCTGCCCGG 918
Oy 321 AlaSerLeuProLysValSerTyrValLysAlaIleAspIleTyrMetAlaValCysLeu 340
Db 919 GCCCTTGGCTTAAGGTGCTCTACGTAGAGCAATCCAGATCTGATGCTGTGTCTG 978
Oy 341 LeuPheValIlePheAlaIleLeuLeuGlyIleTyrAlaIleAlaIleAsnPheValSerArgGlnHis 360
Db 979 CTCTTTGGTTCGCTGCTGCTGCTGAGATGCTGCCATTAATTTGTTTCTGTCAGCAT 1038
Oy 361 LysGlnPheIleArgLeuArgArgArgGlnArgArgGlnArgLeuGlnAspIleIle 380
Db 1039 AAAGAAATTCATACGACTTCGAGAGAGCGAGCGCCAAAGCTTGAGAGAAATATCATC 1098
Oy 381 GlnGlnSerArgPheTyrPheArgGlyTyrGlyLeuGlyHisCysLeuGlnAlaArgAsp 400
Db 1099 CAAGAAGTCTTCTTCTATTCCTGCGCATATGCTTGCGCCACTGCTCCAGCAAGAGAT 1158
Oy 401 GlyLeuProMetGlyLeuSerGlyIleTyrSerProGlnProProAlaProLeuLeuArg 420
Db 1159 GGAGCTCCAAAGGAGGTTTGGCATTTATAGTCCCAACCTCCAGCCCTCTTAAGC 1218
Oy 421 GlnGlyGluThrThrArgGlyLeuTyrValAsp 431
Db 1219 GAAGGAGAACCCACGCGAAGACTCTACGTGAC 1251

RESULT 3
AAQ25722
ID AAQ25722 standard; DNA; 657 BP.
XX
AC AAQ25722;
XX
DE 08-DEC-1992 (first entry)
XX
DE GABA-A receptor beta-subunit.
XX
KW Expression plasmid; gamma-aminobutyric acid; OmpF; glycine; ss.
XX
OS Synthetic.
XX
PN JP04144683-A.
XX
PD 19-MAY-1992.
XX
PE 05-OCT-1990; 90JP-0267743.
XX
PR 05-OCT-1990; 90JP-0267743.
XX
PA (MITU) MITSUBISHI KASEI CORP.
XX
WI: 1992-214122/26.
XX
DR
XX
PT Prep. of N-terminal extracellular site protein - by culturing
XX
PT E.coli transformed by a plasmid comprising the tac promoter, ribosome
XX
PT binding site, etc.
XX
PS Claim 1; Page 2; 12pp; Japanese.
XX
CC The sequence given is the gamma-aminobutyric acid (GABA) A receptor
XX
CC beta-subunit. This sequence is used in an expression plasmid operably
```

```
CC linked to the tac promoter, a ribosome binding sequence, the E. coli
CC outer membrane protein OmpF signal peptide coding sequence and a
CC sequence coding for the N-terminal extracellular site of either the
CC gamma-aminobutyric acid (GABA) A receptor alpha-subunit or the glycine
CC receptor alpha-subunit. This expression plasmid can be used to
CC transform E. coli to produce an N-terminal extracellular site protein
CC of ion channel direct binding type receptor.
XX
SQ Sequence 657 BP; 171 A; 185 C; 156 G; 145 T; 0 other:

Alignment Scores:
Pred. No.: 1,86e-104 Length: 657
Score: 984.50 Matches: 182
Percent Similarity: 87.34% Conservative: 18
Best Local Similarity: 79.48% Mismatches: 14
Query Match: 43.72% Indels: 15
DB: 13 Gaps: 2

US-10-075-846-4 (1-431) x AAQ25722 (1-657)
Oy 40 GlnProMetSerProSerAspPheLeuAspLysLeuMetGlyArgThrSerGlyTyrAsp 59
Db 16 AAGCTATGTCACCCCTGGACTTCTCGATAGCTTATGGGAAGACTTCTGGGTATGAT 75
Oy 60 AlaArgIleArgProAsnPheLysGlyProProValAsnValThrCysAsnIlePheIle 79
Db 76 GCCAGATACAGACCCCACTTTAAAGTCTCTTGAAGCGATGAGATCCACATCTTCATC 135
Oy 80 AsnSerPheSerSerValThrLysThrThrMetAspPyrArgValAsnValPheLeuArg 99
Db 136 AACAGCTTGGTTTATGCGCCGAGACACACCTGGACATACAGGATCAATCTTCTCAGG 195
Oy 100 GlnGlnTrpAsnAspProArgLeuSerTyrArgGlyIleTyrProAspAspSerLeuAspLeu 119
Db 196 CAGCAGTGGAGACGCCCCCGCTCGCTACCATGATATATCTGACGACTCTGGACCTT 255
Oy 120 AspProSerMetLeuAspSerIleTyrLysProAspLeuPhePheAlaAsnGlyLysGly 139
Db 256 GACCATTCACATGTTGGATTCATCTGATGAGAGCTGCTGTTCTTGGCAATGACAAAGGG 315
Oy 140 AlaAsnPheHisGlnValIleThrTrpAspAsnLysLeuLeuArgIlePheLysAsnGlyAsn 159
Db 316 GCCCATTCCACGAGATCACCGACGACACACAGCTGCTGACATATCTCCCGAACGGCAC 375
Oy 160 ValLeuTyrSerIleArgThrLeuIleLeuSerCysLeuMetAspLeuLysAsnPhe 179
Db 376 GTCCTTACACATCAGATATCACCCTGACTGCTGCGCCATGACGACCTGAAGAATTTTC 435
Oy 180 PrometAspIleGlnThrCysThrMetGlnLeuGlnSerSerIleLeuCysSerPro 199
Db 436 CCGATGGACGTACAGACATGATCATGCAACTGGAA----- 471
Oy 200 LeuProSerLeuSerLeuSerValGlyTyrThrMetLysAspLeuValPheGluTrpLeu 219
Db 472 -----AGCTTGGTTATACATGAAACGACCTCATCTTGAAGTGC--- 510
Oy 220 GluAspAlaProAlaValAlaValAlaGlnGlyLeuThrLeuProGlnPheIleLeuArg 239
Db 511 GAAGAGCAAGAGAGCTGTGACAGTGGACAGATGAGACTGACCGCTCAGTTATCTCAGAG 570
Oy 240 AspGluLysAspLeuGlyCysCysThrLysHisIleTyrAsnThrGlyLysPheThrCysIle 259
Db 571 GAAGAGAAAGATCTGAGATACTGACCAAGCATACACAGAGTAATAATTCACCTGCATT 630
Oy 260 GluValLysPheHisLeuGluArgGln 268
Db 631 GAGGCCGATTCACCTGGAAGCGCAG 657

RESULT 4
ABI99254
ID ABI99254 standard; cDNA; 2404 BP.
XX
AC ABI99254;
```



XX	DE	Nucleotide sequence of the human ataxia gene.	
XX	KW	Human; ataxia: gene therapy; ss.	
XX	OS	Homo sapiens.	
XX	PH	Key	Location/Qualifiers
FT	FT	exon	29850..29921
FT	FT	intron	29922..33025
FT	FT	exon	33026..33155
FT	FT	intron	33156..33444
FT	FT	exon	33445..33514
FT	FT	intron	33513..33751
FT	FT	exon	33752..33975
FT	FT	intron	33976..34114
FT	FT	exon	34115..34195
FT	FT	intron	34196..35759
FT	FT	exon	35760..35901
FT	FT	intron	35902..38781
FT	FT	exon	38782..38996
XX	PN	W0200058461-A1.	
XX	PD	05-OCT-2000.	
XX	PF	23-MAR-2000; 2000WO-EP02600.	
XX	PR	26-MAR-1999; 99EP-0106343.	
XX	PA	(RAPP/) RAPPOLD-HOERBRAND G.	
XX	P1	Rappold-Hoerbrand G;	
XX	DR	WPI; 2000-656166/63.	
XX	PT	Novel nucleic acid sequence encoding human ataxia protein for screening compounds useful for treating disorders relating to mutations in ataxia gene -	
XX	PS	Claim 6; Page 22-44; 47pp; English.	
CC	CC	The present sequence represents the human ataxia gene. The ataxia protein and polynucleotides are useful for diagnosing and treating disorders related to ataxia. Ataxia gene sequences are useful in gene therapy, and as diagnostic tools or reagents for identifying and characterizing genetic defect involved in the disorders and diseases related to ataxia.	
CC	CC		
XX	S0	Sequence 39796 BP; 10630 A; 9337 C; 9355 G; 10474 T; 0 other:	
XX	Alignment Scores:		
XX	Pred. No.:	1.03e-75	Length: 39796
XX	Score:	761.00	Matches: 230
XX	Percent Similarity:	24.12%	Conservative: 3
XX	Best Local Similarity:	33.81%	Mismatches: 2
XX	Query Match:	33.79%	Indels: 732
XX	DB:	21	Gaps: 4

OY	21	ValLeuLeuMrpValAlaLeuAlaLysGluGluValLysSerGlyThrLysGlySerGln	40
Db	33015	CTCTGTGTCAGAGGGGCTTGGCAAAAGAGAAAGTAAATCTGGAAACCAAGGGGTCCTCAG	33074
OY	41	ProMetSerProSerAspPheLeuAspLysLeuMetGlyArgThrSerGlyTyrAspAla	60
Db	33075	CCCATGTGCCCTCTGTGATTTCTTAGAACAACTTATGGGGGGAACATCTGATATGATGCC	33134
OY	61	ArgLLeuMrpProAsnPhe-Lys-----	67
Db	33135	AGGATTCGGGCCCAATTTTAAAGTAAGAAATCTTCATCTATAAAACCTCCTCCGCC	33194
OY	67	-----	67
Db	33195	ACCACCTTTGGCAGCACACACTAGCCAGCCACCCCTATTGCTTCCCTTAAGAAAGAGGC	33254
OY	67	-----	67
Db	33255	TGGAGTTCAGAGGCTGGGGCTGGTTCTTACAGCTCAGAAAGAGACCTCTCACCCTATTGT	33314
OY	67	-----	67
Db	33315	CCATGGCCCTGTGTATCTGGCAGCTGATTTTCCACAGAGTGCCCTGGCTGCAGCTCAGG	33374
OY	67	-----	67
Db	33375	GAAGAGTTCTGGAAAGCTGCTCCAGCCTCTGGAATGCCCTGCCAATGGCCGTGTCTTT	33434
OY	68	-----GlyProProValAsnValIleProAsnIlePheIleAsnSerPheSerG	84
Db	33435	GCTACCTCAGGGCCACCCTGAAAGCTGACCTCAACATCTTCATATAACAGATTTCAGCTC	33494
OY	84	rValThrLysThrThrMet-----	90
Db	33495	CATCACCAAGACGCACAT -GTAAGGAGATTCCCTGCTCCCACTCCAGCCTAGTGA	33553
OY	90	-----	90
Db	33554	GTGGAGGAGGCCACACGATACAGAGCTGACATCTTTGCCATATCAGCCCAAGAAAGCT	33613
OY	90	-----	90
Db	33614	TCCTTTCAGATGAAAAATGCCACTACTGCCCCCTAGATGTGTTCCCAACATTCCTCTCAT	33673
OY	90	-----	90
Db	33674	GGCCCTCTGCCCCCTACAGGCTTGCTGGGCCCCCTGGGAATGGCAATGTTTCTGAAGGCC	33733
OY	91	-----AspTyrIleArgValAsnValPheLeuArgGluIleTyrAsnAs	104
Db	33734	CATATCTGCACCTCCAGGACTACGCGGGGAATGTCTTCTCCGCAACAGGGGATGA	33793
OY	104	pProArgLeuSerTyrArgLysTyrProAspAspSerLeuAspLeuAspProSerMetLe	124
Db	33794	CCACAGCCTGTCCTACACGAGAAATATCCTGTATGACTCTCTGAGACCTGCATCCCATGCT	33853
OY	124	uAspSerLleTyrLysProAspLeuPhePheIleAsnGluLysGlyAlaAsnPheHisG	144
Db	33854	GGACTCTATCTGGAGACCAACCTCTTCTTGTGTAATGAAAAAGGGGCCAACTTCCATGA	33913
OY	144	uValThrThrAspAsnLysLeuLeuArgIlePheLysAsnGlyAsnValLeuTyrSerI	164
Db	33914	GGTACCAACGACAAACAGTTACTGCCCATCTTCAAGAAATGGGAATGTGCTGTACAGCAT	33973
OY	164	e-----	164
Db	33974	CAGGTGCACCGGCTGATACCAAGAGAGATTGCTTAAAGGGAGAAATTTGGATGAAGA	34033
OY	164	-----	164
Db	34034	CTGAGGGGTGGAGAGAGGTTCCCTTGCACCTGTGATGGGCCCCAGTAAGCCGATGTACT	34093

```

OY 165 -----ArgLeuThrLeuLeuSerCysLeuMetAspLeuLysa 178
Db 34094 TCTTCTACTGTCCTCCACTGAGCTGACCTCATTTTGTCTGCTGATGAGACCTCAGAGA 34153
OY 178 snPherometaSpIlegInThrcysThrmcInleuLuser ----- 192
Db 34154 ACTTCCCATGGACATCCAGACCTGCACATGACGCTTGAAGCTGTATATA 34213
OY 192 ----- 192
Db 34214 GAGTCACAGAGACCTGGAATAACAGCCAGTTAACCCGACTCATGACATCAGACAGAG 34273
OY 192 ----- 192
Db 34274 AAAAGGCCACTCAGAGTTGGGAGAACTTACTCAAGACACACCAGCAATCTTGAC 34333
OY 192 ----- 192
Db 34334 AGATGAGACTAGAAACAGGGGCTCCTAACCCCTAGCCAGTGTGTTTTCATCCACAG 34393
OY 192 ----- 192
Db 34394 TTAATTGCTGATCTGGCTGTTCCATGGGCTGGGGAAGGGAAGCATCAGCTTACGA 34453
OY 192 ----- 192
Db 34454 CAAATGTGTTGTGCTGTTTGGGGGCGATTTGAGCGACATCTCTGCCCAACGTA 34513
OY 192 ----- 192
Db 34514 TGTCCCTCTCACAGTGCCTGAGTAGTTCACCTTACCTCTTGAATTAAGACATTG 34573
OY 192 ----- 192
Db 34574 AAGATCCCTATAAAATGCCAGAGTGGGAGGCTGGGCGACTGCTGGAGCCCTGATTCC 34633
OY 192 ----- 192
Db 34634 CCACAAACATAGATACTTTGTAGAGAGTAGAGATGACAGTTCTAGGCACACAGAG 34693
OY 192 ----- 192
Db 34694 TGCCTGAATGAAGAGGTAGCCAGAGGCTCTCTGTGCGAGGGTGCACGCTGAGAGG 34753
OY 192 ----- 192
Db 34754 GCTGCCCAAGTGCAGAGAACCAAGATGACAAGTTGGGTCACTTATTTCCACTC 34813
OY 192 ----- 192
Db 34814 CCTCTCATCTCTCTGGGTCAGAGAGACTGTAGATCTGGAGCCATCCACTAGATGTCA 34873
OY 192 ----- 192
Db 34874 CAGAGAAGCCTTATAAGGCTGTTCCAGCATATAAACTGTATGATAGATGTGTC 34933
OY 192 ----- 192
Db 34934 CTATGTGTGTCAGCTCTTTAGAGTTCTTAAACCACTTTACATCCAGGCTCAATATCCC 34993
OY 192 ----- 192
Db 34994 TTCATGCAACGCTGAATAGTGGGAGAGCAGCTACTATATATCCATTTTACAGAGGG 35053
OY 192 ----- 192
Db 35054 GAAGCGACTTACCCAAAGTCACAGAGCTACTGTGGGTCAAGGCCCAACTGGATGTCTG 35113
OY 192 ----- 192
Db 35114 TGAGCCTTGCTGTTTCTGCTCCCATCTCTACCTGACATGTGCAAAATATGTCATATACGT 35173
OY 192 ----- 192

```

```

Db 35174 GGCATGAGCCCTCATGAGACTGCTATGTATGTGAATACTTGACATGTTGACACATTAG 35233
OY 192 ----- 192
Db 35234 CTTTAAATTTGCCCACTCAAGGCCCTGGCTAGATCTATGTATGACATTTGTCCAT 35293
OY 192 ----- 192
Db 35294 GCATGACTGGCAGGATATAAATAGCTGTCTCTCCCTGAGGTGATGTGGCAGAGTCTG 35353
OY 192 ----- 192
Db 35354 GCTCTGAGGCTAGCTCAGCTTAAAGACATGTGTGGTACACATGTTACTCCCTCGGCT 35413
OY 192 ----- 192
Db 35414 TGAGATATGGCAGGAGGCTCCAGCTGTGTTCTCTTACATCTTTCCAAACCTCTTGACCT 35473
OY 192 ----- 192
Db 35474 GCTTCTAGTCTGTGCAATAATTCTCTTTGGGCTTTTTCAGCTTGAGAACAGGGT 35533
OY 192 ----- 192
Db 35534 AGGAGAGAGTCTGTATCTTCTGATTTACCTGCCACTCAGAGATGGCCCATGAGAGA 35593
OY 192 ----- 192
Db 35594 ACTGACCAAGTCTCTCTCCAGAAAGTTTCTTCCATATGACATGAACTCTCTAA 35653
OY 192 ----- 192
Db 35654 ACCCAGACCTTACTCTCCTACCTGTTCTCTCGCTTCGTATCTTACCCCTACCC 35713
OY 193 -----SerSerIleuCysSerProLeuProSerLeuSerLeuValGlyThr 211
Db 35714 ACCGACATCCATACCTCTGACGCCCTCTGCACTCTGTCTCAGTTGAGTTGACCA 35773
OY 211 eLysAspLeuValIleuGluThrLeuGluAspIleuValIleuValIleuGly 231
Db 35774 TGAAGACCTCGTGTGTTGAGTGGCTGAGACATCTCTGCTGTCCAGTGGCTGAGAGG 35833
OY 231 eutHrLeuProGlnPheIleuArgAspGluLysAspLeuGlyCysThrLysHis 251
Db 35834 TGACTCTGCCCAAGTTATCTTGGGGATGAGAAGATCTAGGCTGTGTACCAACGACT 35893
OY 251 yrasnThrGlyLys 255
Db 35894 ACAACACAGGTAA 35907

RESULT 6
AAD21397
ID AAD21397 standard; cDNA: 2400 BP.
XX
AC AAD21397:
XX
DT 28-JAN-2002 (first entry)
XX
DE R. sanguineus glutamate-gated chloride channel 1 cDNA clone, T32.
XX
KW Brown dog tick; glutamate-gated chloride channel; gluc11;
KW gluc12; crop protection; insecticide; nematocide; acaricide;
KW clone T32; ss.
XX
OS Rhipicephalus sanguineus.
XX
FH Key Location/Qualifiers
FT CDS 617..2170
FT FT /tag= a
FT FT /product= "R. sanguineus gluc11 protein, T32"
FT FT /note= "this region is specifically referred in claim 19"
XX

```





Db 1292 CAAGATCGGCGAGTGCAGATCAGATCACCACCCGCC 1327  
RESULT 8  
AAD26937  
ID AAD26937 standard; DNA; 1609 BP.  
XX AAD26937;  
AC AAD26937;  
XX 09-APR-2002 (first entry)  
XX Heliothis sp. HEG3E(4)-2 plasmid DNA.  
DE Heliothis sp. HEG3E(4)-2 plasmid DNA.  
XX Lepidopteran glutamate-gated chloride channel; insecticide;  
KM HEG3E(4)-2 plasmid; cyclic; circular; ds.  
XX Heliothis sp.  
OS  
FH Key Location/Qualifiers  
FT misc\_feature 1  
FT /tag= a  
FT /note= "Base C is present in the sequence shown in  
the sequence listing"  
XX  
XX US6329174-B1.  
XX 11-DEC-2001.  
XX 13-JUN-2000; 2000US-0592891.  
XX 13-JUN-2000; 2000US-0592891.  
XX (AVET ) AVENTIS CROPS SCIENCE SA.  
XX Wang XM, Sarda XG, Tomalski MD, Wingate VPM.  
XX WPI: 2002-121133/16.  
XX New nucleic acid encoding lepidopteran chloride channel, useful for  
screening agents for insecticidal activity -  
XX Example 1: Column 9-10; 18pp; English.  
XX The invention relates to nucleic acid encoding lepidopteran  
glutamate-gated chloride channel. Glutamate-gated chloride channels  
are a family of ligand-gated chloride channels unique to invertebrates.  
The DNA of the invention is used for recombinant production of  
lepidopteran glutamate-gated chloride channel and this is used,  
optionally in the form of membrane preparations or recombinant cells,  
in specific-binding or functional assays for identifying potential  
insecticides. The present sequence is Heliothis sp. HEG3E(4)-2 plasmid  
DNA used in the exemplification of the invention.  
CC Note: This sequence SEQ.ID.NO.11 is stated to be similar to the  
sequence shown in sequence listing of the specification. However  
these sequences differ.  
XX  
XX Sequence 1609 BP: 411 A; 410 C; 393 G; 395 T; 0 other;  
SQ  
Alignment Scores:  
Pred. No.: 1.75e-76 Length: 1609  
Score: 749.00 Matches: 158  
Percent Similarity: 61.568 Conservative: 55  
Best Local Similarity: 45.668 Mismatches: 109  
Query Match: 33.268 Indels: 24  
DB: 24 Gaps: 7  
US-10-075-846-4 (1-431) x AAD26937 (1-1609)  
QY 31 GluValIysSerGlyThrIysGlySerGlnProMetSerProSerAspPheLeuAspLys 50  
ID AAD21378 standard; DNA; 1197 BP.  
XX 211 GAATGATGAAAGCGTGGGAAAGATCACTTTCGAGAGAAAGAGAGATCTGTGATGAG 270  
QY 51 LeuMetGlyArgThrSerGlyTyrAspAlaArgIleArgPro-----AsnPheLys 67

Db 271 ATCCTGGGC---CCCGGAGAGTACGACGCCAGCATCAGACCTCCGGGATCAACGGCAGCT 327  
QY 68 GlyProProValAsnValThnCyAsnIlePheIleAsnIlePheSerSerValThrLys 87  
Db 328 GATGGCCAGCGGAGTGGAGCGTCAATATATTTGTCCGAAGTATATCAAGATCGATGAC 387  
QY 88 ThrThrMetAspTyrArgValAsnValPheLeuArgGlnIntrPasnAspProArgLeu 107  
Db 388 GTCACATGGATGATCTCCGTACATTAACGTTTCGGAGACATGCTTGAAGTGAACGGCTC 447  
QY 108 SerTyrArgGlnTyrProAspAspSerLeuAspLeuAspProSerMetLeuAspSerIle 127  
Db 448 AAATTCATATATCTGGAGGTCGCCCTCAATATACCTGACACTGATGAAGCCACAGAGTC 507  
QY 128 TrpLysProAspLeuPhePheAlaAsnGlnLysGlyLysAsnIleGlnValThr 147  
Db 508 TGGATGCCCTGATCTTCTCTCCACAGAGAGAGAGTCAATTCACAAATCATCATG 567  
QY 148 AspAsnLysLeuLeuArgIlePheLysAsnGlyAsnValLeuTyrSerIleArgLeuThr 167  
Db 568 CCGAAGCGTATACATCGCATCTTCCCAAGCGCAAGTGGCTGATACATCGAATCTCC 627  
QY 168 LeuIleLeuSerCysLeuMetAspLeuLysAsnPheProMetAspIleGlnThrCysThr 187  
Db 628 CTGACGCTCTGCTGCCCATGAACTCAAGTTCATCCCTCGATAGACAGACCTGCTCG 687  
QY 188 MetGlnLeuGlnUserSerSerIleLeuCysSerProLeuProSerLeuSerLeuSerVal 207  
Db 688 CTCAGGATG-----GCTAGTTAT 705  
QY 208 GlyTyrThrMetLysAspLeuValPheGlnTrpLeuGlnAspAlaProAlaValGlnVal 227  
Db 706 GTTGAACACACAGACAGTACTAGTGTCTCTATGAAAGAGAGCGACCG---GTGCAGGTG 762  
QY 228 AlaGlnGlyLeuThrLeuProGlnPheIleLeuArgAspGlnLysAspLeu---GlyCys 246  
Db 763 GTGAAGAACTTACACCGCTCGCTGTCACGCTG-----GAGAAAGTTCCTCAGTACTAC 816  
QY 247 CysThrIlyshIstYrAsnThrGlyLysPheThrCysIleGlnGlyLysPheIleGln 266  
Db 817 TGCACAGCTAAGACTATACCGGTGAATACAGTTCCTGAGAGGTAGACCTGCTTCAAA 876  
QY 267 ArgGlnMetGlyTyrTyrLeuIleGlnMetYrIleProSerLeuLeuIleValIleLeu 286  
Db 877 CGCAGTTCACGTACTACTACTGATTCACATTCCTCCGTGCTGATGCTGCTGCTGCTGCTG 936  
QY 287 SerTyrValSerPheTyrPheIleAsnMetAspAlaIleAlaProAlaArgValGlyLeu 306  
Db 937 TCTGGGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 996  
QY 307 ThrThrValLeuThrMetThrThrGlnSerSerGlySerArgAlaSerLeuProLysVal 326  
Db 997 ACGACTTACTTAAATGCGCAGCCACGACGTCGACGATCAACGGTCCCTACACCGCGTG 1056  
QY 327 SerTyrValLysAlaIleAspIleTyrMetAlaValCysLeuLeuPheValPheAlaAla 346  
Db 1057 TCCCTACAGAAAGCCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1116  
QY 347 LeuLeuGlnTyrAlaAlaIleAsnPheValSerArg-----GlnIlyLysGlnPheIle 364  
Db 1117 CTACTAGAGTGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1176  
QY 365 ArgLeuArgArgArgGln 370  
Db 1177 AAGAAAGCAGACGCGAG 1194  
RESULT 9  
AAD21378  
ID AAD21378 standard; DNA; 1197 BP.  
XX AAD21378;  
XX





XX R. sanguineus GluCl DNA probe to clone GABA-gated chloride channel DNA.  
DE Gamma-aminobutyric acid; GABA; tick infestation; mite; antiparasitic;  
KW GABA-gated chloride channel; recombinant expression; domestic animal;  
KW GluCl DNA; probe; ss.  
OS Rhipicephalus sanguineus.  
PN WO200174884-A1.  
XX  
XX 11-OCT-2001.  
XX  
XX 28-MAR-2001; 2001WO-US09955.  
XX  
XX 31-MAR-2000; 2000US-193791P.  
XX  
XX (MERI ) MERCK & CO INC.  
XX  
XX Zheng Y, Cully D, Ludmerer S;  
PI  
DR WPI: 2002-010778/01.  
XX  
XX New polypeptide useful for preventing or treating tick infestation, in  
PT humans, dogs, cattle, horses, deer, or other wild or domesticated  
PT animals, comprises the Dermacentor variabilis gamma-aminobutyric acid  
PT (GABA)-gated chloride channel -  
XX  
XX  
XX Example 1: Page 23-24; 59pp; English.  
XX  
XX The invention relates to gamma-aminobutyric acid (GABA)-gated chloride  
CC channels and their corresponding nucleic acid molecules. GABA-gated  
CC chloride channel proteins and DNA's are useful for preventing and  
CC treating tick infestation, particularly in humans, dogs, cattle, horses,  
CC deer, or other wild or domesticated animals. The nucleic acids are useful  
CC as hybridisation probes or Polymerase Chain Reaction primers for  
CC identifying the presence of Dermacentor variabilis GABA-gated chloride  
CC channel nucleic acid or DNA's encoding for a GABA receptor. The nucleic  
CC acids are also useful for the recombinant expression of D. variabilis  
CC GABA-gated chloride channel proteins. GABA-gated chloride channel  
CC proteins exert toxic effects on other ticks or related parasites such as  
CC mites. The present sequence is Rhipicephalus sanguineus GluCl DNA probe  
CC which is used for cloning Dermacentor variabilis GABA-gated chloride  
CC channel DNA.  
XX  
XX  
XX Sequence 1197 BP: 268 A; 358 C; 311 G; 260 T; 0 other:  
SO  
Alignment Scores:  
Pred. No.: 6,29e-76 Length: 1197  
Score: 742.50 Matches: 157  
Percent Similarity: 61.06% Conservative: 61  
Best Local Similarity: 43.98% Mismatches: 110  
Query Match: 32.97% Indels: 29  
DB: 24 Gaps: 8  
US-10-075-846-4 (1-431) x AMD22082 (1-1197)  
OY 48 LeuAspSerIleTyrPProValAsnValThrCysAsnIlePheIleAsnSerPheSer 64  
DB 7 TTGGACAGCATCATGTGGCCAG--GTCGTATATACCTCAGATCCGGCCATGGGAATT 63  
OY 65 ---AsnPhelysGlyProProValAsnValThrCysAsnIlePheIleAsnSerPheSer 83  
DB 64 AACAAACACAGAGCGCGGCTCTT--GTACCGCTTAACTCTTTGTAAGAATATCGGC 120  
OY 84 SerValThrLysThrThrMetAspTyrArgValAsnValPheLeuArgGlnGlnTyrAsn 103  
DB 121 AGAATTGATGACGTCACCATGAGTACACAGTGCAGAAAGACAGTTCAAGACAGCGGG 180  
OY 104 AspProArgLeuSerTyrArgGluTyrProAspSerSerLeuAspLeuAspProSerMet 123  
DB 181 GACGAGAGACTTCAGTACGACGACTTGGCGGCCAGGTTCCTACCTGACGCTACCGAA 240

OY 124 LeuAspSerIleTyrPProValAsnValThrCysAsnIlePheIleAsnSerPheSer 143  
DB 241 CCGGACAGCATTTTGAAGCCGCGACTCTTTTCTCCAAAGAGAGGACACTTCAC 300  
OY 144 GluValThrThrAspAsnLysLeuLeuArgIlePheLysAsnGlyAsnValLeuTyrSer 163  
DB 301 AACATCATCATGCCACAGTCTTACGCAATACATCCACAGCGGCGACTCTTCAC 360  
OY 164 IleArgLeuThrLysIleLeuSerCysLeuMetAspLeuLysAsnPheProMetAsp1le 183  
DB 361 ATCAGAAATATCTTGTCTCTTCAATGTCATGACGAGAAATTTTATCTTGGATTA 420  
OY 184 GlnThrCysThrMetGlnLeuGluSerSerIleLeuCysSerProLeuProSerLeu 203  
DB 421 CAAATCTCTCTATC-----GTC 438  
OY 204 SerLeuSerValGlyTyrThrMetLysAspLeuValPheGluTyrPLeuGluAspAlaPro 223  
DB 439 ATGTGAGCTATGCGTATACACAGAGACCTGGTGTTCATGAAAGAGGCGATCT 498  
OY 224 AlaValGlnValAlaGluGlyLeuThrLeuProGlnPheIleLeu---ArgAspGlyLys 242  
DB 499 ---GTACAGTCACAAATAATCTCCACTTGCACAGCTTTCACAGCGTTCACAAAC 555  
OY 243 AspLeuGlyCysCysThrLysHisTyrAsnThrLysPheThrCysIleGluValLys 262  
DB 556 GACTAC-----TGCACAGTCGAGCACACCTGCGGTGACACTGCTGCGCGTGCAC 609  
OY 263 PheHisLeuGluArgGlnMetGlyTyrTyrLeuIleGlnMetTyrIleProSerLeuLeu 282  
DB 610 CTGGTGTTCACAGCGGAGATTACGACTACTGATCTCCAGATCTACATCCCTGCTGCATG 669  
OY 283 IleValIleLeuSerTyrPProValSerPheThrPheLysMetAspAlaAlaProAlaArgVal 302  
DB 670 CTGGTCATCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 729  
OY 303 GlyLeuGlyIleThrThrValLeuThrMetThrThrGlnSerSerGlySerArgAlaSer 322  
DB 730 TCGCTGGCGGCACACCTCTCCATCCATGCGCAGATATGCGCATCAACGCTCG 789  
OY 323 LeuProLysValSerTyrValLysAlaIleAspIleThrPheMetAlaValCysLeuLeuPhe 342  
DB 790 CTGCTCTCCGTTTCTTACACCAAGCCATTTACGTGTGAGACGGCGTCTGTCTACCTTC 849  
OY 343 ValPheAlaLeuLeuGluTyrAlaAlaIleAsnPheValSerArgGlnHisLysGlu 362  
DB 850 GTATTGCGGCGGCTCTGAGTTCGCCCTGCTCACTACGCTGCGGCTGATTCACGC 909  
OY 363 PheIleArgLeuArgArgArgGln-----ArgLeuGluGlu 377  
DB 910 CGGACAGACATGCAAGAACAGAACAGAAATGGAGAGCTGAGACCGCCCTGCTGACTCG 969  
OY 378 AspIleIleGlnGluSerArgPheTyrPheArgGlyTyrGlyLeuGlyHis 394  
DB 970 GACCACCTCGAAGAGAGCGGCCACACAGCTTCCCATGAGCGCGCTGTGTCAC 1020  
RESULT 11  
AAD21395  
ID AAD21395 standard; CDNA: 2138 BP.  
XX  
XX AAD21395;  
AC  
XX 28-JAN-2002 (first entry)  
XX  
DE R. sanguineus glutamate-gated chloride channel 1 cDNA clone, T12.  
XX  
XX Brown dog tick; glutamate-gated chloride channel; GluCl1;  
KW GluCl2, crop protection; insecticide; nematocide; acaricide;  
KW clone T12; ss.  
XX  
XX Rhipicephalus sanguineus.  
OS  
XX  
FH key Location/Qualifiers



FT /note= "This region is specifically referred in claim 13"  
 XX WO200174838-A1.  
 XX 11-OCT-2001.  
 XX 28-MAR-2001; 2001WO-US09905.  
 XX 31-MAR-2000; 2000US-193934P.  
 XX (MERI ) MERCK & CO INC.  
 XX Warmke JW, Yang Y, Cully DF, Hamelin MJ;  
 PI WPI; 2001-662963/76.  
 DR P-PSDB; AAE13038.  
 XX  
 PT Novel L-glutamate-gated chloride channel proteins from Rhipicephalus  
 PT sanguineus for identifying compounds which modulate the channel  
 PT proteins, which are useful as insecticides, antheimintics and  
 PT acaricides -  
 XX  
 PS Claim 12; Fig 3; 89pp; English.  
 XX  
 CC The invention relates to Rhipicephalus sanguineus (brown dog tick)  
 CC L-glutamate-gated chloride channel proteins (Gluc1 and Gluc12)  
 CC and nucleic acid molecules encoding such proteins. Gluc1 channel  
 CC proteins are useful for identifying modulators. The compounds  
 CC identified as modulators are useful for insecticidal, mitacidal  
 CC and/or nematocidal treatment for use in animal and human health  
 CC and/or crop protection. The compounds are also useful in screening  
 CC for and selecting compounds active against parasitic invertebrate  
 CC species relevant to animal and human health, including worms,  
 CC fleas, ticks, mites and lice. Heterologous cell lines expressing  
 CC functional Gluc1 and Gluc12 channel functional forms are useful  
 CC for establishing functional or binding assays to identify novel  
 CC Gluc1 channel modulators. The present sequence is R. sanguineus  
 CC Gluc1 cDNA clone, 782.  
 XX  
 SQ Sequence 2289 BP; 509 A; 727 C; 609 G; 444 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 1 728-75 Length: 2289  
 Score: 742.50 Matches: 157  
 Percent Similarity: 61.06% Conservative: 61  
 Best Local Similarity: 43.98% Mismatches: 110  
 Query Match: 32.97% Indels: 29  
 DB: 22 Gaps: 8  
 US-10-075-846-4 (1-431) x AAD21396 (1-2289)  
 QY 48 LeuAspLysLeuMetGlyArgThrSerGlyTrpAspAlaArgLeuArgPro----- 64  
 DB 625 TTGGACACATCATGTGGCCAG---GGTGTATATGCTCAGAGATCCGGCCATGGGATTT 681  
 QY 65 ---AsnPhelysGlyProProValAsnValThrCysAsnIlePheIleAsnSerPheSer 83  
 DB 682 AACACACAGACGAGCGCGCTCTT---GTACGCGTTACACTGTGTAAGAAGTATCGCG 738  
 QY 84 SerValThrLysThrThMetAspPyrArgValAsnValPheLeuArgGlnGlnTrpAsn 103  
 DB 739 AGAATTGATGACGTCCACATGAGTACAGTCAAAAGAGCTTCAGAGAGAGCGCGG 798  
 QY 104 AspProArgLeuSerTyrArgGlyLysTrpAspSerSerLeuAspProSerMet 123  
 DB 799 GACGAGACACTTCATGACGACACTTGGCGCCAGGTTCCGTACTACGCTCACCGAA 858  
 QY 124 LeuAspSerIleTyrPlyProAspLeuPhePheLeuAsnGlyLysGlyAlaAsnPheHis 143  
 DB 859 CCGGACAGACTTTGGAGACCGGACTTTTCTTCAACGAGAAAGAGGACACTTCAC 918  
 QY 144 GluValThrThrAspAsnLysLeuLeuArgIlePheLysAsnGlyAsnValLeuTyrSer 163

DB 919 AACATCATCATGCCAACGCTCTTACGATACATCCCAAGCGGAGCTTCTTACGC 978  
 QY 164 IleArgLeuThrIleLeuSerCysLeuMetAspLeuLysAsnPheProMetAspIle 183  
 DB 979 ATCAGAAATATCTTGCTGGTCTTTCATGTCGATGAAACCGAAATTTTATCTTGGATAA 1038  
 QY 184 GlnThrCysThrMetGlnLeuGluSerSerIleLeuCysSerProLeuProSerLeu 203  
 DB 1039 CAAATCTCTCTATC-----GTC 1056  
 QY 204 SerLeuSerValGlyTyrThrMetLysAspLeuValPheGluTrpLeuGluAspAlaPro 223  
 DB 1057 ATGGTACCTATGGGTATACACAGAGGACTCGGTGTCTTATGTAAGAGGCGCATCT 1116  
 QY 224 AlaValGlnValAlaGluGlyLeuThrLeuProGlnPheIleLeu---ArgAspGluLys 242  
 DB 1117 ---GTACAGTCCACAAAATCTCCACTTGGCAGCTTTCACGCTGGAAGGTTTCAAAAC 1173  
 QY 243 AspLeuGlyCysCysThrLysHisTyrAsnThrGlyLysPheThrCysIleGluValLys 262  
 DB 1174 GACTAC-----TGACACAGTCGAGCAACAGCTGGAGTACAGCTGCTCGCGTGAC 1227  
 QY 263 PheHisLeuGluAlaArgGlnMetGlyTyrTyrIleuIleGlnMetTyrIleProSerLeu 282  
 DB 1228 CTGTGTTCACAGCGGAGTTCAGCTACTACCTGATCCAGATCTACATCCCTGTCGATG 1287  
 QY 283 IleValIleLeuSerTyrPheValSerPheThrIleAsnMetAspAlaAlaProAlaArgVal 302  
 DB 1288 CTGGCATCTGCTCTGGGTGTCTGTGCTGCTCACACCCACCTGATGCCGCGCGATG 1347  
 QY 303 GlyLeuGlyIleThrThrValLeuThrMetThrThrGlnSerSerGlySerArgAlaSer 322  
 DB 1348 TCGCTGGCGGTCACACACCTGCTCAGATGACGACAGATATGGGATGACGATCAAGCTCG 1407  
 QY 323 LeuProLysValSerTyrValLysAlaIleAspIleThrMetAlaValCysLeuPhe 342  
 DB 1408 CTGCTCCGCTTCTTACACCAAGGCAATGACGCTGAGGAGCGGCTGTCTGATCCCTC 1467  
 QY 343 ValPheAlaIleLeuGluGlyTyrAlaIleAsnPheValSerArgGlnHisLysGlu 362  
 DB 1468 GTATTCGGCGGCTCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1527  
 QY 363 PheIleArgLeuAlaArgArgGlnArgGln-----ArgLeuGlu 377  
 DB 1528 CGCGAAGACATGACAGAACAGAGAGGAAATGGAGCTCGAGGCCGCCCTGGACTCG 1587  
 QY 378 AspIleIleGlnGluSerArgPheThrArgGlyTyrGlyLeuGlyHis 394  
 DB 1588 GACCACTGAGAGAGCGCGCCACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1638  
 RESULT 13  
 AAD22072  
 ID AAD22072 standard; DNA; 1614 BP.  
 XX  
 AC AAD22072;  
 XX  
 DT 12-FEB-2002 (first entry)  
 XX  
 DE Dermacentor variabilis clone 5 GABA-gated chloride channel DNA.  
 XX  
 KW Gamma-aminobutyric acid; GABA; tick infestation; mite; antiparasitic;  
 KM GABA-gated chloride channel; recombinant expression; domestic animal; ds.  
 XX  
 OS Dermacentor variabilis.  
 XX  
 FH key Location/Qualifiers  
 FT CDS 1..1614  
 FT /tag= a  
 FT /product= "GABA-gated chloride channel protein"  
 PN WO200174884-A1.  
 XX 11-OCT-2001.

PF	28-MAR-2001: 200IMD-US09955.	
XX		
XX	31-MAR-2000: 2000US-193791P.	
XX		
PA	(MERI ) MERCK & CO INC.	
XX		
PI	Zheng Y, Cully D, Luderer S;	
XX		
DR	WPI: 2002-010778/01.	
XX	P-PSDB: AAE13314.	
PT	New polypeptide useful for preventing or treating tick infestation, in	
PT	humans, dogs, cattle, horses, deer, or other wild or domesticated	
PT	animals, comprises the Dermacentor variabilis gamma-aminobutyric acid	
PT	(GABA)-gated chloride channel -	
XX		
PS	Claim 8; Fig 3: 59pp; English.	
XX		
CC	The invention relates to gamma-aminobutyric acid (GABA)-gated chloride	
CC	channels and their corresponding nucleic acid molecules. GABA-gated	
CC	chloride channel proteins and DNA's are useful for preventing and	
CC	treating tick infestation, particularly in humans, dogs, cattle, horses,	
CC	deer, or other wild or domesticated animals. The nucleic acids are useful	
CC	as hybridisation probes or Polymerase Chain Reaction primers for	
CC	identifying the presence of Dermacentor variabilis GABA-gated chloride	
CC	channel nucleic acid or DNA's encoding for a GABA receptor. The nucleic	
CC	acids are also useful for the recombinant expression of D. variabilis	
CC	GABA-gated chloride channel proteins. GABA-gated chloride channel	
CC	proteins exert toxic effects on other ticks or related parasites such as	
CC	miles. The present sequence is Dermacentor variabilis clone 5 GABA-gated	
CC	chloride channel DNA.	
XX		
SO	Sequence 1614 BP; 362 A; 475 C; 429 G; 348 T; 0 other:	
Alignment Scores:		
Pred. NO.:	3,81e-75	Length: 1614
Score:	737.50	Matches: 164
Percent Similarity:	58.948	Conservative: 70
Best Local Similarity:	41.318	Mismatches: 107
Query Match:	32.758	Indels: 57
DB:	24	Gaps: 13
US-10-075-846-4 (1-431) x AAD22072 (1-1614)		
OY	55 ThSerglYTYrAspAlarGlleArpPsoAsnPhelySgIyPrOpProAlaSnValThr	74
DB	154 ACACGtGGgGtACACAGAGGGGTGAAGGCCAAATTATTTGGCGGCGTCCAGTGGAGTTGGC	213
OY	75 CyAsnIllePhelLeaSnSerPheSerSerValThrlYsTrhTmMetAspTyArgVal	94
DB	214 GTCACTATGCAGTTTACAGCATTAAGTACAGTCTGTGAAGTACAAATGGACTTACTTCT	273
OY	95 AsnValPhelLeuArgSgInglTrpAsnSprOArGleUSeSerTYrArgLuTYrProAsp	114
DB	274 GACTTCTAATTTCCGCAATCGTGGCGGACGACGACTCTCGTTCCAGAAAAGCCAGAC	333
OY	115 ---AspSerLeuAspLeuAspProSerMetLeuAspSerIleTrpIySProAspLeuPhe	133
DB	334 CTCGACAGCATGCACTGTGGCGCTGAAGTGGCCGAGAGATCTGGTACCCGACACCTTC	393
OY	134 PheAlaAsnGlUySgIyAlaAsnPhelSgIuValThrTrpAspAsnLyLeuLeuArg	153
DB	394 TTGCGCAACAGAAAGGCCCTTACTTTCATGCGGCGCACACAGCCCAACATTTCTTCGCC	453
OY	154 IlePhelYsAnGlyAsnValLeuTYrSerIleArgLeuThrLeuIleLeuSerCysLeu	173
DB	454 ATCGGCTCGCGAGAGAGGTTTTCGCAAGTATTCACATGACGGTGACGCGCGCTGCCA	513
OY	174 MetAspLeuLyAsnPherProMetAspIleGlnTrhYcYstTrhMetGlnLeuGlnSerSer	193
DB	514 ATGGATCTTCGATTACTTCGATATGGACACAAAGCGTCACTATATAGATATGAA-----	567

Oy	194	SerTleLeuCysSerProleuProSerLeuSerLeuSerValGlyIrrThrMetLysasp	213
Db	568	-----AGCTTTGGTTATACCAAGAAGAAC	591
Oy	214	LeuValAPheGLuTrpLeuGLuAspAlaProAlaValAGluValAGluLeuThrLeu	233
Db	592	ATCCCGCTACCGGTGGTCGACCGGTACACCCTCCGTCGCATGCCAAGAGAGTAGAGTTG	651
Oy	234	ProGlnPhe-----IleLeuArgASPGLuValaspleuGLycyscYThrLys	249
Db	652	CCGCAAGTTCAAAGTGCTCGGTGCACGCGCAAAGGAAGAGATTGCCCTTAACGACAGA	711
Oy	250	HISTrysAntHcGLyLysPheThrcysIleGLuValLysPheHisLencLuargLmet	269
Db	712	AACACTCCTC-----CGCCTGGTATCT---GAATACGGCTTC-----GCCCCCTCCAG	756
Oy	270	GlyTYrTYrLeuIleGlnMeTYrIleProSerLeuLeuIleValIleLeuSerTPval	289
Db	757	GGCTACTACCTATATCCAGATCTACATCCCGCGCGGATTATCTGGTTATTTCCTGGCTC	816
Oy	290	SerpHeTrpIleasnMetaspAlaalaProAlaargValIGlyLeuGLylIeThrTrval	309
Db	817	TGCTTTTTGGCTCACCGGTACCGCTAGTCCAGCTCGGCTCGCGCTCGCATCACCGTg	876
Oy	310	LeuthrMetThrThcInsersergILysArgAlaSerLeuProLysValserTYrval	329
Db	877	CTCAGAGATGACACACTCATGTCTCAGTACCAACGACCGCTGCCCAAAATATCTCTACGTC	936
Oy	330	LysAlaIleaspIletrpMetAlaValLysLeuLeuPheValPheAlaIleLeuLeuGLu	349
Db	937	AAGACTATTCAGCTCTACCTGGGGCACATGTCTTGTAATGGTGTGTACCGCGCTCTGGAG	966
Oy	350	TYrAlaAlaIleAsnPheValserArgGlnHisLysGluPheIleArgLeuArgARG---	368
Db	997	TTTTTTTTT-----:-----ATCACCATGAGAAAACC	104
Oy	369	-----ArgGlnArgArgGlnArgLeuGLuArgLeuGLuAspIle	379
Db	1045	CGCTGTACAGCAGCTGGCAAAACTTGCAGAGACAACACAGCAGAGATG-CGCCCGGCTTC	110
Oy	380	IleGlnGLuserrArg-----PheTYrPheArgGLYTyrcIlyLeuGLyHisCysLeu	396
Db	1104	TTCCAACGAGCCAAAGCTGTGAGCCCTTGTACGACGATCCTGAGGTATG--CATTGT--	115
Oy	397	GlnAlaArgASPGLyPrometGLugLyI-----SergIlyIlyrSerPro	412
Db	1158	-----CAAGCAGGTGCGTTCGTGTGCGGTTTCTCTGCGCGGTGCATC-----	1200
Oy	413	GlnProProAlaproLeuLeuArgGLuGLyGluThrThrArgLysLeuTYr	429
Db	1203	-----CNAAGACAACCGAGGGAAGACACACACCAACCGGATTTPAC	1241
<b>RESULT 14</b>			
AAD22070	ID	AAD22070 standard; DNA; 1614 BP.	
XX	AAD22070;		
XX	12-FEB-2002	(first entry)	
XX	Derma	centor variabilis clone 8 GABA-gated chloride channel DNA.	
XX	DE	Gamma-aminobutyric acid; GABA; tick infestation; mite; antiparasitic;	
XX	KM	GABA-gated chloride channel; recombinant expression; domestic animal; ds.	
XX	OS	Derma	centor variabilis.
XX	FH	Key	Location/Qualifiers
XX	FT	CDS	1..1614
XX	FT	/product=	"GABA-gated chloride channel protein".
XX	PN	WO200174884-A1.	

XX 11-OCT-2001.  
 PD 28-MAR-2001; 2001WO-US09955.  
 PF 31-MAR-2000; 2000US-193791P.  
 PR (MERI ) MERCK & CO INC.  
 PA zheng Y, Cully D, Luderer S.  
 PI WPI, 2002-010778/01.  
 XX P-PSDB; AAE13312.  
 DR New polypeptide useful for preventing or treating tick infestation, in  
 PT humans, dogs, cattle, horses, deer, or other wild or domesticated  
 PT animals, comprises the Dermacentor variabilis gamma-aminobutyric acid  
 PT (GABA)-gated chloride channel -  
 PS Claim 8; Fig 3; 59pp; English.  
 CC The invention relates to gamma-aminobutyric acid (GABA)-gated chloride  
 CC channels and their corresponding nucleic acid molecules. GABA-gated  
 CC chloride channel proteins and DNA's are useful for preventing and  
 CC treating tick infestation, particularly in humans, dogs, cattle, horses,  
 CC deer, or other wild or domesticated animals. The nucleic acids are useful  
 CC as hybridisation probes or Polymerase Chain Reaction primers for  
 CC identifying the presence of Dermacentor variabilis GABA-gated chloride  
 CC channel nucleic acid or DNA's encoding for a GABA receptor. The nucleic  
 CC acids are also useful for the recombinant expression of D. variabilis  
 CC GABA-gated chloride channel proteins. GABA-gated chloride channel  
 CC proteins exert toxic effects on other ticks or related parasites such as  
 CC mites. The present sequence is Dermacentor variabilis clone 8 GABA-gated  
 CC chloride channel DNA.  
 XX Sequence 1614 BP; 369 A; 475 C; 422 G; 348 T; 0 other:  
 SO  
 Alignment Scores:  
 Pred. No.: 4,366-75 Length: 1614  
 Score: 737.00 Matches: 164  
 Percent Similarity: 58.78% Conservative: 67  
 Best Local Similarity: 41.73% Mismatches: 114  
 Query Match: 32.73% Indels: 49  
 DB: 24 Gaps: 11  
 US-10-075-846-4 (1-431) x AAD22070 (1-1614)  
 QY 55 ThrsergLYTYRAspAlaArgLeuArgProAsnPhelyGlyProProValAsnValThr 74  
 DB 154 ACACGTGGGTACGACAGAGAGGTGAGGCCAAATTTATGGCGGTTCCAGTGAAGTTGGC 213  
 QY 75 CysAsnIlePheLeuArgGlnGlnTrpAsnAspProAlaArgLeuSerTyArgGluTrpProAsp 94  
 DB 214 GTCACATATGACATATATACGACATATAGTACAGTCTCTGAAGTCAAAATGACATTACTTCT 273  
 QY 95 AsnValPheLeuArgGlnGlnTrpAsnAspProAlaArgLeuSerTyArgGluTrpProAsp 114  
 DB 274 GACTTCTATTTCGCGCAATCGCGGAGCGACGATCTCGTCCAGAAAAAGCCACAGC 333  
 QY 115 ---AspSerLeuAspLeuAspProSerMetLeuAspSerIleTrpLysProAspLeuPhe 133  
 DB 334 CTCGAGACATGACTGTGGCGCCCTGAAGTGGCCGAGAGATCTGGTACCCGACACTTC 393  
 QY 134 PheAlaAsnGluLysGlyAlaAsnPhelIsgluValThrThrAspAsnLysLeuLeuArg 153  
 DB 394 TTCGCCACAGACAGAGCGCCCTACTTTCATGCGGACACACCCCAACACTTCTCCGCC 453  
 QY 154 IlePheLysAsnGlyAsnValLeuTySerIleArgLeuThrIleLeuSerCysLeu 173  
 DB 454 ATCGCTCCGAGAGAGGTTTTCGCAATTCACATGACGAGCGTGCAGTCCGCCA 513  
 QY 174 MetAspLeuLysAsnPhelProMetAspIleGlnTrpCysThrMetGlnLeuGlnSerSer 193  
 |||||||: |||||||: |||

DB 514 ATGCATCTCAGATACTTCCGATGGACAGACAGCGTCACTATAGATAGAA----- 567  
 QY 194 SerIleLeuCysSerProLeuProSerLeuSerValGlyTyThrMetLysAsp 213  
 DB 568 -----ACCTTGCTTATACATGAAAGAC 591  
 QY 214 LeuValPheGluTrpLeuGluAspAlaProAlaValAlaGluValGlyLeuThrLeu 233  
 DB 592 ATCCGCTACCGGTGTCGCGAGCTGACACCTCCGTCACATCCGCAACAGAGTAGAGTTG 651  
 QY 234 ProGlnPhe-----IleLeuArgAspGluLysAspLeuGlyCysCysThrLys 249  
 DB 652 CCGGAGTTCAAGTCTCCTCGTACCTCCAAAAGCAAGAGGTTGCCCTAACGACAGGA 711  
 QY 250 HisTyraAsnThrGlyLysPheThrCysIleGluValLysPheHisLeuGlnArgGlnMet 269  
 DB 712 AACCTACTCC-----CGCCTGTATGT---GAATATACGGTTC-----GCCCGCTCCATG 756  
 QY 270 GlyTyTrpLeuIleGlnMetTyrlleProSerLeuLeuIleValIleLeuSerTrpVal 289  
 DB 757 GGCTACTACGTGATCCATGATCCCGCGGCGGATGATCGTGTATTTCTCGGCTC 816  
 QY 290 SerPheTrpIleAsnMetAspAlaAlaProAlaArgValGlyLeuGlyIleThrThrVal 309  
 DB 817 TCTTTTGGCTCCACCGCTGACGTAGTCCAGTCCGCGGCTCGCGCTCACACCGCTG 876  
 QY 310 LeuThrMetThrThrGlnSerSerGlySerArgAlaSerLeuProLysValSerTrpVal 329  
 DB 877 CTCACGATGACACACTATGCTCCAGTACCAAGCAGCGGCTGCCAAATATCTTACCTC 936  
 QY 330 LysAlaIleAspIleTrpMetAlaValGlyLeuLeuPheValPheAlaIleLeuGlu 349  
 DB 937 AAGATATACGACGTCTACTGCGCACATGTTTGTGAATAGTGTTTACCGCGCTCGAG 996  
 QY 350 TyrAlaIleAlaIleAsnPhelValSerArgGlnHisLysGluPheIleArgLeuArgArg 369  
 DB 997 TACGCGCGGTAGATATCTCGGCAAG-----AGAATCACCATGAGAAAACCCCGTGT 1050  
 QY 370 GlnArgArgGlnArgLeuGlnLysPheIleIleGlnLysSerArgPheTyrlleArgGly 389  
 DB 1051 CAGCAGCTGCGCAAACTTGCACAGCAACA-CAGCGAGAGATGCGC-----CGCGGC 1100  
 QY 390 Tyr-----GlyLeuGly-----HisCysLeu 396  
 DB 1101 TTCTTCCAACGAGCCCAAGCTCTGACCCCTGTGATCCAGTCCGTAAGTATCCATGTCAA 1160  
 QY 397 GlnAlaArgAspGlyGlyProMetGluLysSerGlyIleTyrlleSerProGlnProAla 416  
 DB 1161 GACGCTCGGTTCCTGTCAGATTGTCTGCTGCGGT-----GGCATC 1202  
 QY 417 ProLeuLeuArgGluGlyGluThrThrArgLysLeuTyrlle 429  
 DB 1203 CCAAGACACACCGAGGAAAGCACACCAACCGGATTTAC 1241  
 RESULT 15  
 AAD22071  
 ID AAD22071 standard; DNA; 1614 BP.  
 XX  
 AC AAD22071;  
 DT 12-FEB-2002 (first entry)  
 XX  
 DE Dermacentor variabilis clone 9 GABA-gated chloride channel DNA.  
 KW Gamma-aminobutyric acid; GABA: tick infestation; mite; antiparasitic;  
 XX GABA-gated chloride channel; recombinant expression; domestic animal; ds.  
 OS Dermacentor variabilis.  
 XX  
 FT Key Location/Qualifiers  
 FT CDS 1..1614  
 FT /tag= a  
 FT /product= "GABA-gated chloride channel protein"

XX MO200174884-A1.  
 XX 11-OCT-2001.  
 XX 28-MAR-2001; 2001MO-US09955.  
 XX 31-MAR-2000; 2000US-193791P.  
 XX (MERI ) MERCK & CO INC.  
 XX Zheng Y, Cully D, Ludmerer S;  
 DR WPI; 2002-010778/01.  
 DR P-PSDB; AAE13313.  
 PT New polypeptide useful for preventing or treating tick infestation, in  
 PT humans, dogs, cattle, horses, deer, or other wild or domesticated  
 PT animals, comprises the Dermacentor variabilis gamma-aminobutyric acid  
 PT (GABA)-gated chloride channel -  
 XX  
 XX Claim 8; Fig 3; 59pp; English.  
 XX  
 CC The invention relates to gamma-aminobutyric acid (GABA)-gated chloride  
 CC channels and their corresponding nucleic acid molecules. GABA-gated  
 CC chloride channel proteins and DNA's are useful for preventing and  
 CC treating tick infestation, particularly in humans, dogs, cattle, horses,  
 CC deer, or other wild or domesticated animals. The nucleic acids are useful  
 CC as hybridisation probes or Polymerase Chain Reaction primers for  
 CC identifying the presence of Dermacentor variabilis GABA-gated chloride  
 CC channel nucleic acid or DNA's encoding for a GABA receptor. The nucleic  
 CC acids are also useful for the recombinant expression of D. variabilis  
 CC GABA-gated chloride channel proteins. GABA-gated chloride channel  
 CC proteins exert toxic effects on other ticks or related parasites such as  
 CC mites. The present sequence is Dermacentor variabilis clone 9 GABA-gated  
 CC chloride channel DNA.  
 CC  
 XX  
 XX Sequence 1614 BP; 371 A; 474 C; 421 G; 348 T; 0 other:  
 SO  
 Alignment Scores:  
 Pred. No.: 1,45e-74 Length: 1614  
 Score: 732.50 Matches: 161  
 Percent Similarity: 57.838 Conservative: 68  
 Best Local Similarity: 40.668 Mismatches: 113  
 Query Match: 32.538 Indels: 55  
 DB: 24 Gaps: 11  
 US-10-075-846-4 (1-431) x AAD22071 (1-1614)  
 QY 55 ThSerGIYTYrAspAlaArgIleArgProAsnPhelYsgLIyProProValAsnValThr 74  
 DB 154 ACACGTTGGTACGACGACGAGGGTGGCCAAATTAATGCGCGCTCCAGAGGAAGTTGGC 213  
 QY 75 CysAsnIlePhelIleAsnSerPheSerSerValThrIlyThrMetAspTYrArgVal 94  
 DB 214 GTCACTATGCAAGATTATCACTAATCAAGTCTCTCAATAACAAAGCTTACTTCT 273  
 QY 95 AsnValPheLeuArgInGIuITrPAsnAspProArgLeuSerTYrArgGIuTYrProAsp 114  
 DB 274 GACTTCTATTCCGGCAATGCTGGCGGACGACGACTCTCTGTCAGAAAGCCCAAGAC 333  
 QY 115 ---AspSerLeuAspLeuAspProSerMetLeuAspSerIleTrPlyAspLeuPhe 133  
 DB 334 CTCGAGAGCATGACTGTGGCGCTGAAGTGGCCGACAGATCTGGTACCCGACACTTC 393  
 QY 134 PheAlaAsnGIuYsgLIuAlaAsnPhelIsgLIuValThrThrAspAsnLysLeuArg 153  
 DB 394 TTGGCCACAGAGAGACGCTACTTTCATGCGGCCACCAACGCTTCTCCCGC 453  
 QY 154 IlePheLysAsnGIuAsnValLeuTYrSerIleArgLeuThrLeuIleLeuSerCysLeu 173  
 DB 454 ATCGGCTCCGAGAGAGAGTTTTCGCAATTTCTGACTGACGCGTACTGCCAGCTGCCA 513

QY 174 MetAspLeuLysAsnPhelProMetAspIleGIuThrCysThrMetGIuLeuGIuSer 193  
 DB 514 ATGATCTCAGATTAATCTCCGATGACAGACAGCGCTCAATTAAGATGAGA----- 567  
 QY 194 SerIleLeuCysSerProLeuProSerLeuSerLeuSerValGIuTYrThrMetLysAsp 213  
 DB 568 -----AGCTTTGGTTATACCAAGAAAGAC 591  
 QY 214 LeuValPheGIuTrPLeuGIuAspAlaProAlaValGIuValIleGIuGIuLeuThrLeu 233  
 DB 592 ATCCGCTACCGGTGTGGACGAGTGCAGCTCCGCTCCGCTCCCAAGAGATGAGTTG 651  
 QY 234 ProGIuPhe-----IleLeuArgAspGIuLysAspLeuGIuCysCysThrLys 249  
 DB 652 CCGCAGTTCAAGGCTCGGTGCAGCTCCAAAGCAAGAGAGGTGGCTTCAACGACAGA 711  
 QY 250 HistTrAsnThrLysPheThrCysIleGIuValLysPheIleAsnLeuGIuArgIleMet 269  
 DB 712 AACTACTCC-----CCCTCGTATGT--GAATACGGTTTC-----GCCCGCTCCATG 756  
 QY 270 GIuTYrTrLeuIleGIuMetTYrIleProSerLeuLeuIleValIleLeuSerTrPVal 289  
 DB 757 GGCTACTACTGATCCAGATCTACATCCCGCGGATGATCGTATTATTCCTGGGTC 816  
 QY 290 SerPheTrPLeuAsnMetAspAlaAlaProAlaArgValGIuLeuGIuIleThrThrVal 309  
 DB 817 TCGTTTGGCTCCACCGTAAAGGTATGCAAGTCCGCTCCGCGTCCGCTCCACACCGTC 876  
 QY 310 LeuThrMetThrThrGIuSerSerGIuSerArgAlaSerLeuProLysValSerTYrVal 329  
 DB 877 CTCACGATGACACACATCATGTCCAGTACCAACGACCGCTGCCAAATATCTTACGTC 936  
 QY 330 LysAlaIleAspIleTrPMetAlaValCysLeuLeuPheValPheAlaAlaLeuGIu 349  
 DB 937 AAGAGTATCCAGCTCTACTGGGACACATGTTTGTAAATGGTGTATACCGGCTCTGGAG 956  
 QY 350 TYrAlaAlaIleAsnPhelValSerArgGIuHisLysGIuPheIleArgLeuArgArg--- 368  
 DB 997 TACGCCCGGTAGATATCTCGCAAGAGA-----ATCACCATGAGGAAGAAC 1044  
 QY 369 -----ArgGIuArgArgGIuArgLeuGIuGIuAspIle 379  
 DB 1045 CCGTGTACAGAGCTGCCAAACTTGCAGACCAACAGGACAGAGATG-CGCCGACGCTTC 1103  
 QY 380 IleGIuGIuSerArgPheTYrPheArgGIuTYrGIuLeuGIu----- 393  
 DB 1104 TTCCACAGACCAAGCTC-----TGAGCCCTTCTACCAAGCTCTGAAGTATC 1151  
 QY 394 HisCysLeuGIuAlaArgAspGIuGIuProMetGIuGIuSerGIuIleTYrSerProGIu 413  
 DB 1152 CATTTGTCACAGCGTGTCTCGTCAAGTTTGTCTCGTGGCGT----- 1196  
 QY 414 ProProAlaProLeuLeuArgGIuGIuGIuTYrThrArgLysLeuTYr 429  
 DB 1197 ---GGCATCCCAAGGACACAGAGGAGGACACCAACCGATTTTAC 1241

Search completed: June 30, 2003, 23:13:41  
 Job time : 383.186 secs



THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - nucleic search, using frame\_p2n model

Run on: June 30, 2003, 23:02:29 ; Search time 91.5995 Seconds  
(without alignments)  
1442.995 Million cell updates/sec

Title: US-10-075-846-4

Perfect score: 2252

Sequence: 1 MTLVLPATLSFLMLTLFPGO.....POPPAPLREGCTTKRLXYD 431

Scoring table: BIOSUM62  
Xgapop 10.0, Ygapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

## Command line parameters:

-MODEL=frame\_p2n\_model -DEV=xip  
-O=/cg2\_1/USPTO.spool/US10075846/runat\_25062003\_163649\_5167/app\_query.fasta.1.782  
-DB=Issued\_Patents\_NA -OPMT=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPEXT=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bloms62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10075846.rcgn.1.1.84 -runat\_25062003\_163649\_5167 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NBS\_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

## Database :

- 1: Issued\_Patents\_NA.\*
- 2: /cg2\_6/ptodata/1/ina/5A.COMB.seq.\*
- 3: /cg2\_6/ptodata/1/ina/5B.COMB.seq.\*
- 4: /cg2\_6/ptodata/1/ina/6A.COMB.seq.\*
- 5: /cg2\_6/ptodata/1/ina/6B.COMB.seq.\*
- 6: /cg2\_6/ptodata/1/ina/backfile1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Query Length	ID	Description
1	750	33.3	4621	4	US-09-592-891A-13
2	749	33.3	1609	4	US-09-592-891A-11
3	731.5	32.5	1368	4	US-09-130-339-1
4	731.5	32.5	3958	1	US-08-435-933-5
5	731.5	32.5	3958	5	PCT-US96-06035-5
6	728.5	32.3	1640	4	US-09-592-891A-12
7	728.5	32.3	1844	4	US-09-002-361-1
8	726	32.2	2066	3	US-08-072-064-7
9	724.5	32.2	1519	4	US-09-002-361-4
10	724	32.1	1657	1	US-08-554-659-1
11	724	32.1	1657	1	US-08-554-659-3
12	720	32.0	2066	3	US-08-072-064-2

13	720	32.0	2066	3	US-08-072-064-3	Sequence 3, Appl1
14	718	31.9	2066	3	US-08-072-064-5	Sequence 5, Appl1
15	717	31.8	1491	1	US-08-137-614A-3	Sequence 1, Appl1
16	717	31.8	1491	2	US-08-768-301-1	Sequence 1, Appl1
17	712	31.6	1866	1	US-08-417-330A-19	Sequence 19, Appl1
18	687.5	30.5	1970	3	US-08-072-064-9	Sequence 9, Appl1
19	687.5	30.5	1970	5	PCT-US92-08558-3	Sequence 3, Appl1
20	687.5	30.5	2059	5	PCT-US92-08558-2	Sequence 2, Appl1
21	683.5	30.4	1917	4	US-09-627-6508-6	Sequence 6, Appl1
22	683.5	30.4	1917	4	US-09-436-063C-6	Sequence 6, Appl1
23	680.5	30.2	1297	4	US-09-130-339-3	Sequence 3, Appl1
24	680.5	30.2	1555	4	US-08-809-802-11	Sequence 11, Appl1
25	673	29.9	1479	2	US-08-249-112-2	Sequence 2, Appl1
26	673	29.9	1479	5	PCT-US95-06556-2	Sequence 2, Appl1
27	670.5	29.8	1542	1	US-08-249-112-1	Sequence 1, Appl1
28	670.5	29.8	1542	5	PCT-US95-06556-1	Sequence 1, Appl1
29	663	29.4	2601	4	US-09-627-6508-10	Sequence 10, Appl1
30	663	29.4	2601	4	US-09-436-063C-10	Sequence 10, Appl1
31	657	29.2	2544	4	US-09-627-6508-4	Sequence 4, Appl1
32	657	29.2	2544	4	US-09-436-063C-4	Sequence 4, Appl1
33	656	29.1	2508	4	US-09-627-6508-8	Sequence 8, Appl1
34	656	29.1	2508	4	US-09-436-063C-8	Sequence 8, Appl1
35	645	28.6	1732	1	US-08-417-330A-17	Sequence 17, Appl1
36	628	27.9	2635	2	US-08-768-301-3	Sequence 3, Appl1
37	625.5	27.8	1650	1	US-08-459-100A-1	Sequence 1, Appl1
38	625.5	27.8	1650	5	PCT-US94-09589-1	Sequence 1, Appl1
39	619.5	27.5	1707	4	US-08-809-802-7	Sequence 7, Appl1
40	618	27.4	2191	1	US-08-417-330A-11	Sequence 11, Appl1
41	610.5	27.1	1638	1	US-08-417-330A-13	Sequence 13, Appl1
42	603.5	26.8	1652	4	US-09-627-6508-2	Sequence 2, Appl1
43	603.5	26.8	1652	4	US-09-436-063C-2	Sequence 2, Appl1
44	599	26.6	2310	1	US-08-417-330A-15	Sequence 15, Appl1
45	529.5	23.5	669	4	US-09-002-361-7	Sequence 7, Appl1

## ALIGNMENTS

RESULT 1  
US-09-592-891A-13  
Sequence 13, Application US/09592891A  
Patent No. 6329174  
GENERAL INFORMATION:  
APPLICANT: Xiao-Zhou Michelle Wang  
APPLICANT: Xavier Georges Sarda  
APPLICANT: Michael David Tomalski  
APPLICANT: Vincent Paul Mary Wingate  
TITLE OF INVENTION: Heliothis Glutamate Receptor  
FILE REFERENCE: A32815 072667.0118  
CURRENT FILING DATE: 2001-01-22  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 13  
LENGTH: 4621  
TYPE: DNA  
ORGANISM: Heliothis virescens  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (764)...(764)  
OTHER INFORMATION: n is a, t, g, or c  
US-09-592-891A-13

## Alignment Scores:

Pred. No.: 1.77e-85 Length: 4621  
Score: 750.00 Matches: 167  
Percent Similarity: 58.42% Conservative: 62  
Best Local Similarity: 42.60% Mismatches: 139  
Query Match: 33.30% Indels: 25  
DB: 4 Gaps: 7

US-10-075-846-4 (1-431) x US-09-592-891A-13 (1-4621)

QY 31 GluVallysserclyThrlyGlySerClnPrometserProserAsphleuAsplys 50

[illegible]

```

Db      1232 CACAGATGCAACACACCCTTTGCATGTAACAAACCTTGTTGGCGGCGCGTGTGGATTC 1229
Ox      405 GlucylSerGlyIleTyrSerProGlnProAla 416
          ::      ::      ::      ::      ::      ::      ::      ::      ::
Db      1292 CAGATGCGGCGAGTGCAGATCCACATCCACCCCGCC 1327
          ::      ::      ::      ::      ::      ::      ::      ::      ::

RESULT 2
US-09-592-891A-11
: Sequence 11, Application US/09592891A
: Patent No. 6329174
: GENERAL INFORMATION:
: APPLICANT: Xiao-zhou Michelle Wang
: APPLICANT: Xavier Georges Sarda
: APPLICANT: Michael David Tomaski
: APPLICANT: Vincent Paul Mary Wingate
: TITLE OF INVENTION: Heliothis Glutamate Receptor
: FILE REFERENCE: A32815 072667.0118
: CURRENT APPLICATION NUMBER: US/09/592.891A
: CURRENT FILING DATE: 2001-01-22
: NUMBER OF SEQ ID NOS: 14
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 11
:
: LENGTH: 1609
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: PCR amplified fragment of Heliothis DNA cloned
: OTHER INFORMATION: Into pCR2.1-TOPO vector (Invitrogen)
: US-09-592-891A-11

```

[illegible]

Db 688 CTCAGCATG-----GCTAGTAT 705  
Qy 208 GATYrThMetLysAspLeuValPheGluTrpLeuGluAspAlaProAlaValGlnVal 227  
Db 706 GATTGACCAACAGACACTTAGTGTCTTGAAGAGGACCGG--GTGAGAGTG 762  
Qy 228 AlaGluGlyLeuThrLeuProGlnPheIleLeuArgAspGluLysAspLeu---GlyCys 246  
Db 763 GTGAAAACCTACACTGCGCTCGCTCAGCGTG-----GAGAAAGTTCCTCACTGACTAC 816  
Qy 247 CysThLysHisTrpAsnThrGlyLysPheThrCysIleGluValLysPheHisLeuGlu 266  
Db 817 TGCACAGTAAGACTATACCGGTGATACAGTGTGCTGAGAGTAGACCTGCTCTTCAA 876  
Qy 267 ArgGlnMetGlyTrpTrpLeuIleGlnMetTrpIleProSerLeuLeuIleValIleLeu 286  
Db 877 CGCGAGTTCAAGTACACCTGATCCAGATCTACATTCCGTGCTGATGCTGATCCTG 936  
Qy 287 SerTrpValSerPheTrpIleAsnMetAspAlaAlaProAlaArgValGlyLeuGlyIle 306  
Db 937 TCCTGGGTGCTCTTCTGCGCGAGACAGAGCTGTGCTGCGAGGTCTCACTAGAGAGTG 996  
Qy 307 ThrThValLeuThrMetThrThrGlnSerSerGlySerArgAlaSerLeuProLysVal 326  
Db 997 ACGACTTACTTACATGCGCGACCCAGTCGTCAGCATCAACCGCTCCCTACACCGGTG 1056  
Qy 327 SerTrpValLysAlaIleAspIleTrpMetAlaValCysLeuLeuPheValPheAlaIle 346  
Db 1057 TCCTTACACGAAACCCATGTATGTCTGACTGGGTATGTCTCACTTCGATTCGAGAGCG 1116  
Qy 347 LeuLeuGluTrpAlaIleAsnPheValSerArg-----GlnHisLysGluPheIle 364  
Db 1117 CTACTAGAGTTGCGCTCGCACTATGCTCTGCTGACATGCACGACGAGAACATG 1176  
Qy 365 ArgLeuArgArgArgGln 370  
Db 1177 AAGAAAGCGAGCGGAG 1194  
RESULT 3  
US-09-130-339-1  
; Sequence 1, Application US/09130339  
; Patent No. 638701  
; GENERAL INFORMATION:  
; APPLICANT: Cully, Doris F.  
; APPLICANT: Paresse, Philip S.  
; APPLICANT: Warmke, Jeffrey W.  
; APPLICANT: Eiter, Adrian  
; APPLICANT: Cohen, Charles J.  
; APPLICANT: Brochu, Richard M.  
; TITLE OF INVENTION: DNA MOLECULES ENCODING CTENOCEPHALIDES  
; TITLE OF INVENTION: FELIS GLUTAMATE GATED CHLORIDE CHANNELS  
; FILE REFERENCE: 20029  
; CURRENT APPLICATION NUMBER: US/09/130, 339  
; CURRENT FILING DATE: 1998-08-06  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 1368  
; TYPE: DNA  
; ORGANISM: ctenocephalides felis  
US-09-130-339-1  
Alignment Scores:  
Pred. No.: 5, 03e-84 Length: 1368  
Score: 731.50 Matches: 157  
Percent Similarity: 59.78% Conservative: 60  
Best Local Similarity: 43.25% Mismatches: 111  
Query Match: 32.48% Indels: 35  
DB: 4 Gaps: 9  
US-10-075-846-4 (1-431) x US-09-130-339-1 (1-1368)

Qy 24 ArgValAlaLeuAlaLysGluGluValLysSerGlyThrLysGlySerGlnProMetSer 43  
Db 79 CGTCTAGAGAGCGGCAAGAAATTTCCAGGCGCCAAAGAAAG----- 120  
Qy 44 ProSerAspPheLeuAspLysLeuMetGlyArgThrSerGlyTrpAspAlaValIleArg 63  
Db 121 -----CAAGCTTGGACCAAAATTTTAAAGC--CCAGGCCATTACGATGCCAAATTAAG 171  
Qy 64 ProAsn-----PheLysGlyProProValAsnValThrcysAsnIlePhe 78  
Db 172 CTTCTGGAGCTCAATGAGACTGAGACGGTCCAGCCGTG---GTACAGCTCAACATCTAT 228  
Qy 79 IleAsnSerPheSerSerValThrLysThrThrMetAspTrpArgValAsnValPheLeu 98  
Db 229 CTGAGATTCATATCAGCGAATAGATGACTACAAATGGAATAGATGCTCCAGTAACTTTC 288  
Qy 99 ArgGlnGlnTrpAsnAspProAlaGluSerTrpArgGluTrpProAspAspSerLeuAsp 118  
Db 289 AGCGAACAATGCGAGGTGAGAGTGAATTTTAAAGACTTTTGAGAGTCTTTTAAATAC 348  
Qy 119 LeuAspProSerMetLeuAspSerIleTrpLysProAspLeuPheAlaAsnGluLys 138  
Db 349 TTACACACTACCGACAGCACTCGTGTATGAGATGCCGATTTGTTTGGAAATGAAG 408  
Qy 139 GlyAlaAsnPheHisGluValThrThrAspAsnLysLeuLeuArgIlePheLysAsnGly 158  
Db 409 GAGGGCCACTTTCACAAACATCATCATCCGACGTGCTGATTTTCTTACCGT 468  
Qy 159 AsnValLeuTrpSerIleArgLeuThrLeuIleLeuSerCysLeuMetAspLeuLysAsn 178  
Db 469 TCCGTACTATACAGCATCAGAGATATCGCTTACCTTGTGGCGTCTATGAATTCGAACATG 528  
Qy 179 PheProMetAspIleGlnThrCysThrMetGlnLeuGlnSerSerIleLeuCysSer 198  
Db 529 TATCCGCTCATAGCGCAGGTGCTCTCTCCGATG----- 564  
Qy 199 ProLeuProSerLeuSerLeuSerValGlyTrpThrMetLysAspLeuValPheGluTrp 218  
Db 565 -----GCCAGTATGTTGGTGGACCAACAAACGATCGTGTTTTCTG 606  
Qy 219 LeuGluAspAlaProAlaValGlnValAlaGluGlyLeuThrLeuProGlnPheIleLeu 238  
Db 607 AAGGAAGTGACCGG---GTGACAGTTGTCAAGAACTTACATCTGCCACAGTTTACCTTG 663  
Qy 239 ArgAspGluLysAspLeu---GlyCysCysThrLysHisTrpAsnThrGlyLysPheThr 257  
Db 664 -----GAGAAAGTTCGTGACCGATTAATTGAACAGCAAAACCAATACCGTGAATACAGT 717  
Qy 258 CysIleGluValLysPheHisLeuGluArgGlnMetGlyTrpTrpLeuIleGlnMetTrp 277  
Db 718 TGCCTGAGAGTGCACCTGCTCTTTAAACGAGAGTTCGACATCACTGATCCAGATCTAC 777  
Qy 278 IleProSerLeuLeuIleValIleLeuSerTrpValSerPheTrpIleAsnMetAspAla 297  
Db 778 ATTCTTGTTCATGTTGCTGATCGTTTCTCGGTGCTGTTGTTGAGCCAGGAGCG 837  
Qy 298 AlaProAlaArgValGlyLeuGlyIleThrThrValLeuThrMetThrGlnSerSer 317  
Db 838 GTTCCGGCCAGAGATCATCTGGGTGACCACTCTCTCAACAGGCCACCGACGACCTCG 897  
Qy 318 GlySerArgAlaSerLeuProLysValSerTrpValLysAlaIleAspIleTrpMetAla 337  
Db 898 GGCATTAAGCGCTCCCGCCGCGCCAGTGTCTTACACAAAGCCATCGACGTGAGCCGA 957  
Qy 338 ValCysLeuLeuPheValPheAlaIleLeuGluTrpAlaIleAsnPheValSer 357  
Db 958 GTCTGCTTACAGTTCGTTCTTGGGGCTTCTCGAAATTCGCCCTCGCACTAGCGCTCC 1017  
Qy 358 Arg-----GlnHisLysGluPheIleArgLeuArgArgGlnArgGlnArgLeu 375  
Db 1018 AGATCCGATATGCACAGGAGAAACATGAAACAAAAGCGCAGGAACTTGAACAAGCAGCC 1077  
Qy 376 GluGluAsp 378

Db 1078 AGCTGAC 1086

RESULT 4  
US-08-435-933-5  
Patent No. 5693492  
Sequence 5, Application US/08435933  
GENERAL INFORMATION:  
APPLICANT: Cully, Doris F.  
APPLICANT: Arena, Joseph P.  
APPLICANT: Pares, Philip S.  
APPLICANT: Liu, Ken K.  
TITLE OF INVENTION: DNA ENCODING GLUTAMATE GATED CHLORIDE  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: John W. Wallen III  
STREET: 126 East Lincoln Avenue  
CITY: Rahway  
STATE: New Jersey  
COUNTRY: US  
ZIP: 07065-0907  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/435,933  
FILING DATE:  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Wallen, III John W.  
REGISTRATION NUMBER: 35,403  
REFERENCE/DOCKET NUMBER: 19264  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 594-4720  
TELEFAX: (908) 594-4720  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3958 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-435-933-5

Alignment Scores:  
Pred. No.: 3,27e-83  
Score: 731.50  
Percent Similarity: 57.96%  
Best Local Similarity: 42.30%  
Query Match: 32.48%  
Length: 3958  
Matches: 162  
Conservative: 60  
Mismatch: 127  
Indels: 35  
Gaps: 8

US-10-075-846-4 (1-431) x US-08-435-933-5 (1-3958)

Qy 3 ThrLeuValProAlaThrLeuSerPheLeuLeuLeuProGlyGlnValLeu 22  
Db 390 ACCCCATTCCACATGGGACGACACTATTTCGGCG-----ATCTTA 437  
Qy 23 LeuArgValAlaLeuAlaLysGluGluValLysSerGlyThrLysGlySerGlnProMet 42  
Db 438 TACTTGCACGCTGCTGCTTCACTACCAATATATGCAAGTAAATTCGAGAA 497  
Qy 43 SerProSerAspPheLeuAspLysLeuMetGlyArgThrSerGlyTyrAspAlaArgLe 62  
Db 498 AAGGAGAAAAAGCTTAGATCAATTTTAGT---CGAGCAATACGAGCCCGAATA 554  
Qy 63 ArgPro-----AsnPhelyGlyProProValAsnValThrCysAsnIlePheIle 79  
Db 555 CGACCACTGTGAATTAATGGCAGCATGTGTCGCCCATAGTCAGCAATCACTATTCGT 614

Qy 80 AsnSerPheSerValThrLysThrMetLysPtyrArgValAsnValPheLeuArg 99  
Db 615 CGCAGTATTATGACATAGTATGATATTAATGAGTAGTACAGTACAGTTAACCTCCGT 674  
Qy 100 GlnGlnTrpAsnAspProArgLeuSerTyrArgGlyTyrProAspAspSerLeuAspLeu 119  
Db 675 GAACGTGACGAGTGAAGCCCTCAAGT--CGACGATATCCAGGCGCCCTAAAGTATCT 733  
Qy 120 AspPro-SerMetLeuAspSerIleTrpLysProAspLeuPhePheAlaAsnGlyLysG 139  
Db 734 GACCTGACGAGCGAAGCCCGCTGTGATCCCGATCTTTCTTCGACGAGAGGA 793  
Qy 139 yAlaAsnPhelyGluValThrTrpAspAsnLysLeuLeuArgIlePheLysAsnGlyLys 159  
Db 794 GGCACACTTCCACAAATCATCATGCGCAATGTATATTCGATCTTCCCAACGATC 853  
Qy 159 nValLeuTyrSerIleArgLeuThrLeuIleLeuSerCysLeuMetLysLeuLysAsnPh 179  
Db 854 TGTGCTATATAGTATACATCTCGCTGACATTCGCGCCCAATGAACCTAAAGCTGTA 913  
Qy 179 eProMetAspIleGlnThrCysThrMetGlnLeuGluSerSerIleLeuCysSerPr 199  
Db 914 TCCCGTGATGACATGATCTCTCTACGAGATC----- 947  
Qy 199 OleuProSerLeuSerLeuSerValGlyTyrThrMetLysAspLeuValPheGluTrpLe 219  
Db 948 -----GCCACTATAGCTGTGAGCACCAACAGCAGTGGTCTTCGTGGA 991  
Qy 219 uGluAspAlaProAlaValAlaGluGlyLeuThrLeuProGlnPheIleLeuArg 239  
Db 992 GGAGGCGCATCC--GTACAGGTGTAAAGAACTTACCTACCTCGCTTACACATG-- 1046  
Qy 239 gAspGluLysAspLeu---GlyCysCysThrLysHisTyrAsnThrGlyLysPheTrpC 258  
Db 1047 ----GAGAACTTCTGACTGATTAAGTAAAGCAACACCGGTGAATACAGTTCG 1102  
Qy 258 sIleGluValLysPheHisLeuGluArgGlnMetGlyTyrTrpLeuIleGlnMetYrI 278  
Db 1103 CCTCAAGTCCATGATCTATTCAGCGAGAAATTCATATTAATTAATTAAT 1162  
Qy 278 eProSerLeuLeuLeuValIleLeuSerTrpValSerPheTrpLysAsnMetLysAla 298  
Db 1163 ACCATGCTGATGTGCTATGATCATGAGGTATCATCTGCTGATCAAGAGCAGT 1222  
Qy 298 aProAlaArgValGlyLeuGlyIleThrThrValLeuThrMetThrTrpGlnSerSerG 318  
Db 1223 ACCGCGGAGTGTACATGGGTGTACACACCTCTGACATGCCACCGACGATCGCG 1282  
Qy 318 ySerArgAlaSerLeuProLysValSerTyrValLysAlaIleAspIleTrpMetAlaVal 338  
Db 1283 CATAAAGCCCTCCCTGCGCCGCTTCTTATACGAAGGCCATGATGTGACAGCGCT 1342  
Qy 338 lCysLeuLeuLeuPheAlaAlaLeuLeuGluTyrAlaAlaIleAsnPhelySerArg 358  
Db 1343 GTGTCTGACGTTCGTTCGCGGCGCTGCTGATTCGCGCTGTGAATCATGATCCCG 1402  
Qy 358 g-----GlnHisLysGluPheIleArgLeuArgArgGlnArg 371  
Db 1403 ATCAGTTCGAAATTAAGCTAATGATGATAGAGAAATATGAAAGAACGCCCGCATCT 1462  
Qy 371 gArgGln 373  
Db 1463 GGAGCAG 1469

RESULT 5  
PCT-US96-06035-5  
Sequence 5, Application PC/TUS9606035  
GENERAL INFORMATION:  
APPLICANT: Cully, Doris F.  
APPLICANT: Arena, Joseph P.  
APPLICANT: Pares, Philip S.  
APPLICANT: Liu, Ken K.  
TITLE OF INVENTION: DNA ENCODING GLUTAMATE GATED CHLORIDE



US-09-592-891A-12

### Alignment Scores:

Pred. No.:	1,69e+83	Length:	1640
Score:	728.50	Matches:	56
Percent Similarity:	61.10%	Conservative:	156
Best local Similarity:	44,964	Mismatches:	110
Query Match:	32,358	Indels:	25
DB:	4	Gaps:	8

US-10-075-846-4 (1-431) X US-09-592-891A-12 (1-1640)

OY	31	GUValuLysSerGlyThrLysGlySerGlnProMetSerProSerAspHeuAspLys	50
		:	:
Db	154	GNATCATGAAAGCGGTGGCAAGATCACTTTCAGAGAGACAGACAGACATCCTGGATCCAG	213
OY	51	LeuMetGlyArgThrSerGlyTyrAspAlaArgIleArgPro-----AsnPhelys	67
Db	214	ATCCCGGGC---CCCGGAGGTAGACGCCAGGATCAGACCCCGGGATTCACAGCGGCAC	270
OY	68	GLY---ProProValAsnValThrCysAsnIlePheIleAsnSerPheSerIleThr	86
		:    :	:
Db	271	GGCTATGGCCACGCTTACCTCCATGTCACATGATATCTACGGTCCATCAGCAAAATAGAT	330
OY	87	LysThrThrMetAspTyrArgValAsnValPheLeuArgGlnIleThrPasnAspProArg	106
		:        :    :	
Db	331	GATTACAAATATGAGATATACCTCCGTCACAAATTAAGCTTTCGGGACCAATGGTTAGTAGAACG	390
OY	107	LeuSerTyrArgGluTyrProAspAspSerLeuAspLeuAspProSerMetLeuAspSer	126
			:
Db	391	CTCAAAATTCATATATCTGGAGGTGCGCTCAAAATCCTACACTGATCAACCAACCAAGA	450
OY	127	IleTrpLysProAspLeuPhePheAlaAsnGluLysGlyAlaAsnPheHisGluValThr	146
			:
Db	451	GTCTGATGAGCCTGATCTATTTCTTCTCCACGAGGAAGGAGGTGATTTCCACAAACATC	510
OY	147	ThrAspAsnLysLeuLeuArgIlePheLysAsnGlyAsnValLeuTyrSerIleArgLeu	166
		:	:
Db	511	ATGCCGAACGTGTACATCCGGATCTTCCCAACGGCAACGCTGTACAGCAATCCCAATC	570
OY	167	ThrLeuIleLeuSerCysLeuMetAspLeuLysAsnPheProMetSerPheIleThrCys	186
			:
Db	571	TCCCTGAGAGCTCTCGTCCGCCCATGAACTCAGATTGTACCCCTGGATTAAGCAGACTGC	630
OY	187	ThrMetGlnLeuGlnSerSerIleLeuCysSerProLeuProSerLeuSerLeuSer	206
		:	
Db	631	TCCGACAGAGATG-----GCTAGT	648
OY	207	ValGlyTyrThrMetLysAspLeuValPheGluTrpLeuGluAspAlaProAlaValGln	226
		:	
Db	649	TATGTTGTGGACACAGACAGCTAGTACTGTTCCATATGGAAGGAAGGCACCG---GTGCAG	705
OY	227	ValAlaGluGlyLeuThrLeuProGlnPheIleLeuArgAspGluLysAspLeu---Gly	245
		:	
Db	706	GTGGGAAAACCTTACACCTGCTCGGTCCAGCGTG-----GAGAAGTTCCTCACCTGAC	759
OY	246	CysCysThrLysHisTyrAsnThrGlyLysPheThrCysIleGluValPheHisLeu	265
		:    :    :	
Db	760	TACTGCACACAGTAAGACTATATACCGGTGATACAGTGTGCTGAAGGTAGACCTGCTTTC	819
OY	266	GluArgGlnMetGlyTyrTyrLeuIleGlnMetTyrIleProSerLeuLeuIleValIle	285
		:	:
Db	820	AAAGCGAGTTCAGTTACTACCTGATGCCAGATCTCAATCCGTCGTCGATGCGTGCATC	879
OY	286	LeuSerTrpValSerPheTrpIleAsnMetAspAlaAlaProAlaIleArgValGlyLeuGly	305
		:	
Db	880	GTGTCTCTGGGTCTCTCTGCTGACCCAGGAGCTGCTCTCGCGAGGCTTCATCATTAGGA	939
OY	306	IleThrThrValLeuThrMetThrThrGlnSerSerGlySerArgIleAspLeuProLys	325
		:	
Db	940	GTGACGACTTACTTACATAGCGACCCAGTGTGTCAGGATCAACCGGTCTCTACACCG	999
OY	326	ValSerTyrValLysAlaIleAspIleTrpMetAlaValCysLeuLeuPheValIleAla	345

Db	1000	GTGTCCTACACGAAGCCATTGATGCTCGGACGTGGATTATCTTCACATTTTGTATTCCGA	1059
Qy	346	AlaLeuLeuGluTyrAlaAlaAlaLeuAspPheValIserArg-----GlnHisLysGluPhe	363
Db	1060	GCGCCTACTAGAGTTTGGCCCTGCCTCAACTATGCGTCTGCTGCATGCACCGAGAGAAC	1119
Qy	364	ILeaArgLeuArgArgArgGln	370
Db	1120	ATGAAGAAAGCGAGACGGGAG	1140

## RESULT 7

US-09-002-361-1  
; Sequence 1, Application US/09002361

```

: GENERAL INFORMATION:
: APPLICANT: Halling, Blak
: TITLE OF INVENTION: Lepidopteran GABA-gated Chloride
: TITLE OF INVENTION: Channels
: NUMBER OF SEQUENCES: 43
: CORRESPONDENCE ADDRESS:
:

```

```

1  COMPUTER READABLE FORM:
2  MEDIUM TYPE: Diskette
3  COMPUTER: IBM Compatible
4  OPERATING SYSTEM: DOS
5  SOFTWARE: FASTEO for Windows Version 2.0
6  CURRENT APPLICATION DATA:
7  APPLICATION NUMBER: US/09/002.361

```

```

;
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:

```

REGISTRATION NUMBER: 29.135  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 609-520-3214  
TELEFAX: 609-520-3259  
TELEX:

```

? INFORMATION FOR SEQ ID NO: 1:
?
? SEQUENCE CHARACTERISTICS:
?
? LENGTH: 1844 base pairs
?
? TYPE: nucleic acid
?
? STRANDEDNESS: single
?
? TOPOLOGY: linear
?
? FEATURE:
?
? NAME/KEY: Coding Sequence
?
? LOCATION: 104...1591
?
? OTHER INFORMATION:
?
```

US-09-002-361-1

Alignment Scores:	
Pred. No.:	2, 08e+83
Score:	728.50
Percent Similarity:	57.77%
Best Local Similarity:	38.83%
Query Match:	32.35%
DB:	4
Gaps:	10
Length:	1844
Matches:	160
Conservative:	78
Mismatches:	119
Indels:	56
Gaps:	10

US-10-075-846-4 (1-431) x US-09-002-361-1 (1-1844)

QY 17 LeuProGlnValLeuLeuArgValAlaLeuAlaLysGluGluValLysSerGlyThr 36  
 ||||| : : : : : |||||  
 Db 162 TTGCGTGGTTACCTCATGCTGCATCCATGCCGCGGAGCGGAGAGAG-----GGGGATGT 215



Oy	37	LysGlySerGlnProMetSerProSerAspPheLeuAspLysLeuMetCylarGlnHisSer	56
Db	216	TT-GGTGAC-----GTCAATATTCACGCCATTGGATTGGATTCGTA-----AGTGA	259
Oy	57	GIYTRAspAlaArgIleArgProAsnPhenylSGlyProProValAsnValThyCysAsn	76
Db	260	ACGTCACGAAAGAGTCAGCGCCAACTATGGAGAGACCCCACTGATGTGGAGACAC	319
Oy	77	IlePheIleAsnSerPheSerSerValThrIlyStrThrMetAspTryArgAsnVal	96
Db	320	ATGTACGTCGCTCCATCAGCTCCTTATCTGAAGTAAATGATTCACCTGATTC	379
Oy	97	PheLeuArgGlnGlnIleTrpAsnAspProArgLeuSerTyArgGluIleProAsp---Asp	115
Db	380	TACTTCACACAAATTTTGGACACAGCCCAAGCTTGGCTTTACAAAAAGACGGGTGGAG	439
Oy	116	SerLeuAspLeuAspProSerMetLeuAspSerIleTryIlyProAspLeuPheAla	135
Db	440	ACTCTGTCGCGCGCTCGGAATTTATTAGAAACATATGGGTACCCGACACCTTCTTGGT	499
Oy	136	AsnGluIlySGlyAlaAsnPhenylSGluValIleThrAspAsnLysLeuLeuArgIlePhe	155
Db	500	AACGAAAAACAGCTTATTTCACATAGCTACTACAGACAGAAATTCATAGCAATTCAT	559
Oy	156	LysAsnGlyAsnValLeuTyIleSerIleArgLeuThrIleLeuSerCysLeuMetAsp	175
Db	560	CATTCGTGATCTATTACTAGAGATTAACAGTACACTATCACGCTTCTTGTCGAGAT	619
Oy	176	LeuLysAsnPhenProMetAspIleGlnIleThyCysThrMetGlnLeuIleuSerSerIle	195
Db	620	TTGCGATTTTCCATGATGACCGCTCAATATTCATATTCGAATTCGAA-----	667
Oy	196	LeuCysSerProLeuProSerLeuSerLeuSerValGlyTyIleThrMetLysAspLeuVal	215
Db	668	-----AGTTTGGCTACACCATCGCGGACATCCGA	697
Oy	216	PheGluTrpPheGlnAspAlaProAlaValGlnValAlaGluIlyLeuThrIleProGln	235
Db	698	TACAGTGGAAATGAGGGGCCAACCTACAGTGGGTGTGTGACGGAAGCTGTTGGCCGCAA	757
Oy	236	PheIleLeuArgAspGluLysAspLeuGly-----CysCysThlyShiStyI	251
Db	758	TTC-----AAGTCTGGGGCCACCGGACGGGCCATGAGATTTCTCT	802
Oy	252	AsnThrGlyLysPheThrCysIleGluValLysPheHisLeuGlnIlyArgGlnMetGlyTyI	271
Db	803	ACGACAGGAACTACTCTGCTGCGATGTGAATTCATTTGTACGTCGATGGGATAC	862
Oy	272	TyrLeuIleGlnMetTyIleProSerLeuLeuIleValIleLeuSerTryValSerPhe	291
Db	863	TATTTAAATTCAGATTTATTTCTCCGCTCGCCCAATGTGATATATCTTGGGTATCATTT	922
Oy	292	TrpIleAsnMetAspAlaIleProAlaArgValGlyLeuGlyIleThrThrValLeuThr	311
Db	923	TGGTTGAATCGAAATGCGACACCTGCAGAGGTATCTAGGTGTCACTGATTTGACG	982
Oy	312	MetThrTrpGlnSerSerGlySerArgAlaSerLeuProLysValSerTyValIlyVala	331
Db	983	ATGACAGCGCTCATGTGCTGTCACAGAAATGGCGCTGCCCAAGATCTCATATGCAAGTGC	104
Oy	332	IleAspIleTrpMetAlaValCysLeuLeuPheValPheAlaIleLeuGluIlyTyAla	351
Db	1043	ATCGATGCTCTAATCTGGGAATGTTTCATGATGCTTGTGCCAGTTTACTAGAAATATGCC	110
Oy	352	AlaIleAsnPhenValSerArgGlnHisLysGluPheIleArgLeuArgIleArgGlnArg	371
Db	1103	ACGGTTGGCTAATGGCTAAA-----AGCATACAGATGACGAAACAA	114
Oy	372	ArgGlnArgLeuGluGluAspIleIleGlnIleuSerArgPheTyIleArgGlyTyGly	391
Db	1145	ACATTTCACGCTGTTCCAAAAAATGGCGCGGACGAGAAATAATGCAATA	119
Oy	392	LeuGluHisCysLeuGlnIleAlaArgAspGlyGlyProMetCyluIlySerGlyIleTyIle	411

```

Db      1193  -----GATGGTCTCCAGGTCAGTGCACCTATCCCCCA 1228
Oy      412  Pro-----GlnProAlaPro 417
          |||
          ::|||
Db      1229  CCGAGGACGACACCTATCTAGGCCACACCACT 1264

RESULT 8
US-08-072-064-7
: Sequence 7, Application US/08072064
: Patent No. 6008046
: GENERAL INFORMATION:
: APPLICANT: FRENCH-CONSTANT, RICHARD H.
: APPLICANT: JACKSON, MEYER B.
: TITLE OF INVENTION: DRUG AND PESTICIDE SCREENING
: NUMBER OF SEQUENCES: 20
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: PETER G. CARROLL
: STREET: 220 Montgomery Street, Suite 2200
: CITY: San Francisco
: STATE: California
: COUNTRY: United States of America
: ZIP: 94104
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/072,064
: FILING DATE: 19930602
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 770,881
: FILING DATE: 04-OCT-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: CARROLL, PETER G.
: REGISTRATION NUMBER: 32,837
: REFERENCE/DOCKET NUMBER: OPID-00574
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415/705-8410
: TELEFAX: 415/397-8338
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2066 base pairs
: TYPE: NUCLEIC ACID
: STRANDEDNESS: single
: TOPOLOGY: unknown
: MOLECULE TYPE: CDNA
US-08-072-064-7

Alignment Scores:
Pred. No.:      5,33e-83      Length:      2066
Score:          726.00        Matches:      156
Percent Similarity: 56.94%    Conservative: 86
Best Local Similarity: 36.71% Mismatches:    143
Query Match:     32.24%      Indels:        40
DBs:             3           Gaps:           9

US-10-075-846-4 (1-431) x US-08-072-064-7 (1-2066)
Oy      5 ValProAlaThrLeuSerPheLeuLeuLeuTrpPheLeuProGlyGlnValLeuLeuArg 24
          ::||| ||| ||| ::|||
          CCGCTGCTAATACATCTG-----CTGGCATTCAAC 178
Db      137 CCGCCCCCGCACAC-----ValIysSerGlyThr 36
          ::||| ||| ::|||
          ATGGCCCTGATGGACAGGAAACGGGCCCAAAACGATCCATACAGTGCACACGGCGACT 238
Oy      25 ValAlaLeuAlaLysGluGlu-----ValIysSerGlyThr 36
          ::||| ||| ::|||
          ATGGCCCTGATGGACAGGAAACGGGCCCAAAACGATCCATACAGTGCACACGGCGACT 238
Oy      37 LysGlySerGlnProMetSerProSerAspPheLeuAspIysLeuMetGlyArgThrSer 56
          ||| ||| ::||| |||
          GCGGCTGGCAGCATCTGGGTGACGTAAACATATCCGCTATTCTGACCTCTTTAGTGTT 298

```



[illegible]

```
US-08-659-1  
Sequence 1, Application US/08554659  
Patent No. 5767261  
GENERAL INFORMATION:  
APPLICANT: Wingate, Vincent  
TITLE OF INVENTION: LEPIDOPTERAN GABA GATED CHLORIDE CHANNEL  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Scully, Scott, Murphy & Presser  
STREET: 400 Garden City Plaza  
CITY: Garden City  
STATE: New York  
COUNTRY: USA  
ZIP: 11530-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/554,659  
FILING DATE:  
CLASSIFICATION: C36  
ATTORNEY/AGENT INFORMATION:  
NAME: Digiglio, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 9732  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 742-4343  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO.: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1657 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 115..1581  
US-08-659-1
```

Alignment Scores:

Pred. No.	6,54e-83	Length:	1657
Percent Similarity:	724.00	Matches:	165
Score:	55.68%	Conservative:	80
Best Local Similarity:	37.50%	Mismatches:	115
Query Match:	32.15%	Indels:	81
DB:	1	Gaps:	12

```
US-10-075-846-4 (1-431) x US-08-659-1 (1-1657)
```

OY    17 leuProGlyglnValleuleuarqyAlAalaLeuAtLySGlucglUvallysEr---gly 35  
     ||||| | : : : : : | | : : : : :  
Db     173 TTGGCTGTACTCATCGTGCAGACATGCCGCCGGAGCCGGAGAGGGGAATGTTGGTG 23  
OY      36 ThrIysGLYSerclnPrmetSerPserAspPhelenuSpLYSleumetcLYArqThr 55  
     ||| ||||||| | : : : : :  
Db      233 ACGCAATAATCATCACCA-----TTTTG-GATTTCCTA-----AGT 26

OY    56 SerGIlyTrspAlarqlrIArqprroasnphelysgLYpproproVALsnVAlThrcys 75  
     ||||| | : : : : : | | : : : : :  
Db      268 GTAAGCAACAACAAAGAAGTAGCGCACTATGGAGACCAGCCAGTGTAGTGGACATC 32

OY    76 AsnllepheleasnsErsPseservAlThrlYsthThmetaspTYfargvAlasn 95  
     ||::: :::: ||| ||:::::::::: ||||| : : : : :  
Db      328 AACATGTAAGTGCCTTCATCAGCTCTTATCTGAAGTCAAALATGATTTTCA CCTGAT 38

OY    96 ValpleuLeuqrGlncInfrrpasnasPPoArqleuSeTyfArgglutrytProasp-- 114  
     ::| ||||| | || ||||||| ::::: ||| ::::: :

Db 388 TTCTACTCAGACAAATTTTGGACAGACCCAGGCTTCTTACAAAAGACAGCGGTGTG 447  
Qy 115 AspsertleuaspleuasProserMetleuasSerletrplysProaspLeuPhePhe 134  
Db 448 GAGACTGTGCGCGGTGGGATTTATTAAGAAACATATGGGACCGACACCTTCTT 507  
Qy 135 AlaasnleuaspleuasPheHisgluValThrThrAspAsnlysLeuLeuArg 154  
Db 508 GTTAACGAAAAACAGCATATTTCCACATAGCTACTACAGCAACGAAATTCATACGAT 567  
Qy 155 PheylasnglyAsnValleuThrSerThrleargleuThrleuLeuSerCysleuMet 174  
Db 568 CATCATTTCTGATCTTACTAGAGATTAAGACTATACACCGCTTCTTGTCCGAT 627  
Qy 175 AspleuasnPhetPheMetAspPheGlnThrCysThrMetGlnleuGluSerSer 194  
Db 628 GATTGGAGATATTTCCGATGGACCGCATATATGCAATATGAAATCGAA----- 678  
Qy 195 IleleucysSerProleuProSerleuSerleuSerValGlyThrThrMetLysAspleu 214  
Db 679 -----AGTTTGGCTACACCATCGGCGACATC 705  
Qy 215 ValPheGluThrPheleuGluAspAlaProAlaValGlnValAlaGluGlyLeuThrPro 234  
Db 706 CGATACAGATGGAATGAGGGGCCCACTCACTGCTGTGTCGAGCGAAGTCTTCTTCCG 765  
Qy 235 GlnPheIleleuArgAspGluLysAspleuGly-----CysCysThrLysHis 250  
Db 766 CAATTC-----ANSGTGTGGCGCCACCGCGCGGCGCATGGAATTTCT 810  
Qy 251 TyrAsnThrGlyLysPheThrCysIleGluValLysPheHisleuGluArgGlnMetCly 270  
Db 811 CTACAGCAGAGAACTACTCTCGTGTGCGCATGTGAATTCATTTAGGCTCGATGGA 870  
Qy 271 TyrThrleuIleGlnMetClyThrleuProSerleuLeuValIleleuSerThrValSer 290  
Db 871 TACTATTAAATGATTAATTCCTGCTGCGCTTAATTCATTAATTCCTTGGGTACA 930  
Qy 291 PheThrPheAsnMetAspAlaAlaProAlaArgValGlyleuGlyIleThrThrValleu 310  
Db 931 TTTTGGTTGAATCAAAATGCGACACGTCGAAAGGTATACAGTGTGCACACACTGTATG 990  
Qy 311 ThrMetThrThrGlnSerSerGlySerArgAlaSerleuProLysValSerThrValLys 330  
Db 991 ACGATGACAGCGCTCATGTCGCCAGAAATGCGGCTCGCCCAAGATCTCATATGTCAG 1050  
Qy 331 AlalleaspleuMetAlaValCysleuLeuPheValPheAlaAlaLeuGlnThr 350  
Db 1051 TCCATCGATGCTATCTGCGAAGCTTTGCTCATGGCTCTTCCACAGCTTACTAGAAATAT 1110  
Qy 351 AlaAlaIleAsnPheValSerArgGlnHisLysGluPheIleArgLeuArgArgGln 370  
Db 1111 GCCACGGTTGGCTATATGGCGAA-----AGCATACAGATAGGAAA 1152  
Qy 371 ArgArg-----GlnArgLeuGluLysP----- 378  
Db 1153 CAAGATTCACACTGCTGTCAAAAATCAAAATAGATGTCCTCCAGGGTACAGTACGCT 1212  
Qy 379 -----IleIleGlnGluSer 383  
Db 1213 ATCCCCCAGGAGGACGACCCCTATCTAGGCCACACCTAGCCGATTAATGCGAGGTT 1272  
Qy 384 ArgPhe-----TyrPheArgGlyLysGlyLeuGlnHisCysleu 396  
Db 1273 CGGTTCAAAAGTCCAGATCCGAAAGCATATTTCTAAAGCGGTACTTTTAAAAACATATC 1332  
Qy 397 GlnAlaArgAspGlyLysPheMetGluGlySerGlyIleThrSerProGlnProAla 416  
Db 1333 AATGGGCTGCGGCGCAGCA-----GGACCTGCTCCACCGCA 1371

RESULT 11  
US-08-554-659-3  
; Sequence 3, Application US/08554659

Patent No. 5767261  
GENERAL INFORMATION:  
APPLICANT: Wingate, Vincent  
TITLE OF INVENTION: LEPIDOPTERAN GABA GATED CHLORIDE CHANNEL  
TITLE OF INVENTION: AND METHODS OF USE THEREOF  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESS: Scully, Scott, Murphy & Presser  
STREET: 400 Garden City Plaza  
CITY: Garden City  
STATE: New York  
COUNTRY: USA  
ZIP: 11530-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/554,659  
FILING DATE:  
CLASSIFICATION: 436  
ATTORNEY/AGENT INFORMATION:  
NAME: Digilio, Frank S  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 9732  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1657 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 115..1581  
US-08-554-659-3  
Alignment Scores:  
Pred. No.: 6,54e-83 Length: 1657  
Score: 724.00 Matches: 165  
Percent Similarity: 55.68% Conservative: 80  
Best Local Similarity: 37.50% Mismatches: 115  
Query Match: 32.15% Indels: 81  
Gaps: 12  
US-10-075-846-4 (1-431) x US-08-554-659-3 (1-1657)  
Qy 17 LeuProGlyGlnValleuLeuArgValAlaLeuAlaLysGluGluValLysSer---Gly 35  
Db 173 TTGGCTGTGTAACCATGACCATGCGGCGGAGCGAGGAGGAGGAGGAGGAGGAGGAGG 232  
Qy 36 ThrLysGlySerGlnProMetSerProSerAspPheleuasplysLeuMetGlyArgThr 55  
Db 233 ACGTCATATCTCGCA-----TTTTC-GATTGCGTA-----AGT 267  
Qy 56 SerGlyTyrAspAlaArgIleArgProAsnPheLysGlyProProAlaAsnValThrCys 75  
Db 268 GTAGCTACGCAAAAGAGTGAAGGCGGACATATGAGAGCGGACGAGTGTGGAGTGC 327  
Qy 76 AsnIlePheIleAsnSerPheSerSerValThrLysThrThrMetAspTyrArgValAsn 95  
Db 328 AACATGTACGCTCTCATCAGCTCCTTATCTGAAGTGAATAATGATTTACACCTGGAT 387  
Qy 96 ValPheleuArgGlnGlnThrPAsnAspProArgleuSerThrIleArgGluThrProAsp--- 114  
Db 388 TTCTACTCAGACAAATTTTGGACAGACCCAGGCTTCTTACAAAAGACAGCGGTGTG 447

```

OY 115 AspSerLeuAspLeuAspProSerMetLeuAspSerIleThrPlyProAspLeuPhe 134
Db 448 GAGACTCTCCGCGCGCGGATTTATAGAAATATGGTACCACACCTTCTT 507
OY 135 AlaAsnGluGlyAlaAsnPheHisGluValThrThrAspAsnLysLeuLeuArgIle 154
Db 508 GTTAACGAAACAGCATATTTCCACATAGCTACACAAACAAATTCATACCAATT 567
OY 155 PheLysAsnGlyAsnValLeuThrSerIleArgLeuThrLeuLeuLeuSerLeuMet 174
Db 568 CATCATCTCTGATCTACTAGAGATTAAGACTACTACTACCCCTCTTGTCCGATG 627
OY 175 AspLeuLysAsnPheProMetAspIleGlnThrCysThrMetGlnLeuGlnSerSer 194
Db 628 GATTTCAGATATTTTCCGATGGAGCCGTCATATTCATATTCGAATTCGAA----- 678
OY 195 IleLeuCysSerProLeuProSerLeuSerLeuSerValGlyThrThrMetLysAspLeu 214
Db 679 -----AGTTTGGCTACACCAATGCGGAGATC 705
OY 215 ValPheGluThrLeuGluAspAlaProAlaValGlnValAlaGluGlyLeuThrLeuPro 234
Db 706 CGATACAGAGGATAGGGGCCCAACTAGTGGTGTGCGAGCAAGTGTTCGCG 765
OY 235 GlnPheIleLeuArgAspGlyLysAspLeuGly-----CysCysThrLysHis 250
Db 766 CAATTC-----AAGTGTCTGGCGCCAGCGGCGCATGGAGATTCT 810
OY 251 TyrAsnThrGlyLysPheThrCysIleGluValLysPheHisLeuGluArgGlnMetGly 270
Db 811 CTYACACAGCAAGAACTACTCTCGTCGCGATGGAATTCATATTGTAGCGCTGATGGA 870
OY 271 TyrThrLeuIleGlnMetThrIleProSerLeuLeuIleValIleLeuSerThrValSer 290
Db 871 TACTATTATTAATTCAGATTATATTCGCGCTGCGCTATGTCATTTATCTGGGATCA 930
OY 291 PheThrIleAsnMetAspAlaAlaProAlaArgValGlyLeuGlyIleThrThrValLeu 310
Db 931 TTTTGGTTGAATGGAATGCACTGACACAGGATGACACTAGGTGTCAACATGTATTG 990
OY 311 ThrMetThrThrGlnSerSerGlySerArgAlaSerLeuProLysValSerThrValLys 330
Db 991 AGCATACACAGCGCTCATGTCGTCACGAATGCGGCTCTGCGCAAGATCTCATATGTCAAG 1050
OY 331 AlaIleAspIleThrPheAlaValLysLeuLeuPheValPheAlaAlaLeuLeuGluThr 350
Db 1051 TCCATGATGTCTATGTGGAACTGTTCGTCATGCTCTCACCAAGTTTACTAGATAT 1110
OY 351 AlaAlaIleAsnPheValSerArgGlnHisLysGluPheIleArgLeuArgArgGln 370
Db 1111 GCCAGGTTGGCTATATGGCGAAA-----AGCATACAGATGAGGAAA 1152
OY 371 ArgArg-----GlnArgLeuGluGluAsp----- 378
Db 1153 CAAGATTCACGTCTGTCAAAAATGCAAAATAGATGTCCTCCAGGTCAGTGAGCT 1212
OY 379 -----IleIleGlnGlnSer 383
Db 1213 ATCCCCCAGCAGACACACCCATCTAGCGCACACCATAGCCGATATTCGAGGTT 1272
OY 384 ArgPhe-----TyrPheArgGlyTyrGlyLeuGlyHisCysLeu 396
Db 1273 CGGTTCAAGTTCACAGATCCGAAGGATATTCGAAAGCGGCTACTTGAAGAACTATC 1332
OY 397 GlnAlaArgAspGlyLysPheMetGluGlySerGlyIleThrSerProGlnProAla 416
Db 1333 AATGGGCTCGGGCGGCAAGCA-----GGACCTGCTCCACCGCA 1371

```

RESULT 12  
 US-08-072-064-2  
 ; Sequence 2, Application US/08072064  
 ; Patent No. 608046  
 ; GENERAL INFORMATION:

```

; APPLICANT: FRENCH-CONSTANT, RICHARD H.
; APPLICANT: JACKSON, MEYER B.
; TITLE OF INVENTION: DRUG AND PESTICIDE SCREENING
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PETER G. CARROLL
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/072,064
; FILING DATE: 19930602
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 770,881
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: CARROLL, PETER G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: OPND-00574
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/705-8410
; TELEFAX: 415/397-8338
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2066 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; US-08-072-064-2

Alignment Scores:
Pred. No.: 3.16e-82 Length: 2066
Score: 720.00 Matches: 155
Percent Similarity: 56.71% Conservative: 86
Best Local Similarity: 36.47% Mismatches: 144
Query Match: 31.97% Indels: 40
Gaps: 9

US-10-075-846-4 (1-431) x US-08-072-064-2 (1-2066)
OY 5 ValProAlaThrLeuSerPheLeuLeuThrProGlyGlnValLeuLeuArg 24
Db 137 CTGCCCCGCGACA---CCGCTGTAACCATCTG-----CTGGCCATCAAC 178
OY 25 ValAlaLeuAlaLysGluGlu-----ValLysSerGlyThr 36
Db 179 ATGGCCCTGATTCACAGAGAAAGCGGCCAACAGATCCATACATGCAAGGCGGACT 238
OY 37 LysGlySerGlnProMetSerProSerAspPheLeuAspLysLeuMetGlyArgThrSer 56
Db 239 GCGCGTGGCAGCATGCTGGGTGAGTAACATATCCGCTATCTCGACACTCTTAACTGT 298
OY 57 GlyTyrAspAlaArgIleArgProAsnPheLysGlyProProAlaAsnValThrCysAsn 76
Db 299 AGTTACGACAAAGAGTAAGACCAATTAAGGTTGCCCCCTGTGGAGTTGGCGTCACA 358
OY 77 IlePheIleAsnSerPheSerSerValThrLysThrThrMetAspTyrArgValAsnVal 96
Db 359 ATGTATGTCTCTAGTATCAATTCGTTTGGAGGTTCTATAGCACTTCATGATTTT 418
OY 97 PheLeuArgGlnIleThrPheAsnAspProArgLeuSerTyrArgGluTyrProAsp---Asp 115
Db 419 TACTTTCGCAATTTTGGACGATCTCGTTTGTAGCGATAGAAAAGACCTGCTGTAGAA 478

```

```

OY 116 SerLeuAspLeuAspProSerMetLeuAspSerIleTrpLysProAspLeuPheAla 135
Db 479 ACACATCGCTGGATCGAGCTTCAATTAAGATATTTGGTACCTGACACCTTTTGTATA 538
OY 136 AsnGlySerGlyAlaAsnPheHisGluValThrThrPaspAsnLysLeuLeuArgIlePhe 155
Db 539 AATGAAAAACAATATTTATTTGACATTGCAACAAACAGTAATGATTCATCGTGTGAT 598
OY 156 LysAsnGlyAsnValLeuTyrSerIleArgLeuThrLeuIleLeuSerCysLeuMetAsp 175
Db 599 CATCTCGATCGAATAACAGAAAGATTAAGATTGACATAACCGCATCGTCCGATGAT 658
OY 176 LeuLysAsnPheProMetAspIleGlnThrCysThrMetGlnLeuGluSerSerIle 195
Db 659 CTACAAATATTTCCCATGATGCGCAGCTGTGCCACATTTGAATGAA----- 706
OY 196 LeuYsSerProLeuProSerLeuSerLeuSerValGlyTyrThrMetLysAspLeuVal 215
Db 707 -----AGCTTGGTTACAGATTCGAGATATCCGA 736
OY 216 PheGluTrpLeuGluAspAlaProAlaValGlnValGluGlyLeuThrLeuProGln 235
Db 737 TATTTCTGGAGAGATGAGCTGAGTAGTGTGGCATGACAGTGAAGTGAAGTGAAGTGAAG 796
OY 236 PheIleLeuArgAspGluLysAspLeuGlyCysCysThrLysHisTyrAsnThrGlyLys 255
Db 797 TTTCGAGTTTGGACACAGGAGCAG---AGGCGAGCAGAAATAACCTAACACAGGAGCAG 853
OY 256 PheThrCysIleGluValLysPheHisLeuGlnValArgGlnMetGlyTyrThrLeuIleGln 275
Db 854 TATTCGCGTTTACGCTGCGAATTCACAGTCTGCTGCTGATGAGGCTGCTGCTGCTGCTGCT 913
OY 276 MetTyrIleProSerLeuLeuIleValIleLeuSerTrpValSerPheTrpIleAsnMet 295
Db 914 ATCTACATACCTCTGACATGATGCTGTATATCATGATGATATTTGGCTCATGCGC 973
OY 296 AspAlaIleProAlaArgValGlyLeuGlyLeuThrThrValLeuThrMetThrGln 315
Db 974 AATCGAACGCGCGCGCTGCGTGGCTGCGTGCACCCGCTGACAAATGACCACTTTCG 1033
OY 316 SerSerGlySerArgAlaSerLeuProLysValSerTyrValLysAlaIleAspIleTrp 335
Db 1034 ATGCGTCACAAATGCGCGCTGCCAAAGATTTCGATGATGATGATGATGATGATGATGAT 1093
OY 336 MetAlaValCysLeuLeuPheValPheAlaIleLeuGluGlyTyrAlaIleAsnPhe 355
Db 1094 CTGGAGACATCTCTGATGATGCTCTTGGCCAGTCAATGACGACGAGTGGCGTAC 1153
OY 356 ValSerArgGlnHisLysGluPheIleArgLeuArgArgGlnArgGlnArgLeu 375
Db 1154 ATGGCAAAA-----CGAATTCAAAATGCGAAAAAAGATTTATGCGC 1195
OY 376 GluIleuAspIleIleGlnGluSerArgPheArgPheArgGlyTyrGlyLeuGlyHisCys 395
Db 1196 ATCCAAAGATAGCGCAACAAAAAGACACAGCTCCAGGAGGAGGAGGAGGAGGAGGAGGAG 1255
OY 396 -----LeuGlnAlaArgAspGlyLysProMetGluGlySerGlyIleTyrSerPro 412
Db 1256 AATCCCAATCCCAATGCAATATGTCGCGAGCC---GGAGAGTGGGCGTTGGAGCCCGC 1312
OY 413 GlnProProAlaPro 417
Db 1313 GGACCCGGAGGAGCC 1327

```

```

CORRESPONDENCE ADDRESS:
ADDRESS: PETER G. CARROLL
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC-compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/072,064
FILING DATE: 19930602
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 770,881
FILING DATE: 04-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: CARROLL, PETER G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: OPHD-00574
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/705-8410
TELEFAX: 415/397-8338
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2066 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: CDNA
US-08-072-064-3
Alignment Scores:
Pred. No.: 3,16e-82 Length: 2066
Score: 720.00 Matches: 155
Percent Similarity: 56.71% Conservative: 86
Best Local Similarity: 36.47% Mismatches: 144
Query Match: 31.97% Indels: 40
DB: 3 Gaps: 9
US-10-075-846-4 (1-431) x US-08-072-064-3 (1-2066)
OY 5 ValProAlaThrLeuSerPheLeuLeuThrPheLeuProGlyGlnValLeuLeuArg 24
Db 137 CTGCCCGGACACA---CCGCTCTTAACCACTTGG-----CTGGCCATCAAC 178
OY 25 ValAlaLeuAlaLysGluGlu-----ValLysSerGlyThr 36
Db 179 ATGGCCCTGATTGACAGAGAAACGCGCCACAAACGATTCATACAGTGAAGCGGCACT 238
OY 37 LysGlySerGlnProMetSerProSerAspPheLeuAspLysLeuMetGlyArgThrSer 56
Db 239 GCGCGTGGACACATCGCGGTGACGTAAACATTCCTATTCCTGACTCCCTTAGAGTT 298
OY 57 GlyTyrAspAlaArgIleArgPheProAsnPheLysGlyProValAsnValThrCysAsn 76
Db 299 AGTTACGCAAAAGATAGAACCAATTAACCGTGTCCCTGTGAGAGTTGGCGTCA 358
OY 77 IlePheIleAsnSerPheSerSerValThrLysThrThrMetAspTyrArgValAsnVal 96
Db 359 ATGTATGTCCTCACTATACAGTTCGTTTGGAAAGTTCATAGCACTTCACATTTGATTTT 418
OY 97 PheLeuArgGlnGlnTrpAsnAspProArgLeuSerTyrArgGlyTyrProAsp---Asp 115
Db 419 TACTTTGCTCAATTTTGGACCGATTCGCTTTTACCGATAGAAACGACCTGTGTAGAA 478
OY 116 SerLeuAspLeuAspProSerMetLeuAspSerIleTrpLysProAspLeuPheAla 135
Db 479 ACACATCGCTGGATCGAGCTTCAATTAAGATATTTGGTACCTGACACCTTTTGTATA 538

```

```
QY 136 AsngluysgLyAlaAsnPhenHisGluValThrThrAspAsnLysLeuLeuArgIlePhe 155
DB 539 AATGAAAAAATCATATTTTCAATGTAACACAGTAATGAAATTCATACGTGAT 598
QY 156 LysAsngLyAsnValLeuTyrSerIleArgIleThrLeuIleLeuSerCysLeuMetAsp 175
DB 599 CATTTGGATCGATTAACAAGAAAGTATAGTACTAATAACCGCATCGCTCCGATGAT 658
QY 176 LeuLysAsnPhenProMetAspIleGlnThrCysThrMetGlnLeuSerSerIle 195
DB 659 CTACAAATTTTCCCATGATGATGCGACGCTGCCACATTGAATGAA----- 706
QY 196 LeuLysSerProLeuProSerLeuSerLeuSerValGlyTyrThrMetLysAsnVal 215
DB 707 -----ACCTTCGGTTACCATGCGAATATCCGA 736
QY 216 PheGluThrPheGluAspAlaProAlaValGlnValAlaGluGlyLeuThrProGln 235
DB 737 TATTTCTGAGAGAGATGAGATGATGTTGCGATGACAGCAGTGCAGTACCCGAG 796
QY 236 PheIleLeuArgAspGluLysAsnLeuGlyCysCysThrLysHisThrAsnGlyLys 255
DB 797 TTCCGAGTTTGGGACACAGGAG---AGGGCGACCGAAATTAACCTAACCGACGAAAC 853
QY 256 PheThrCysIleGluValLysPheHisLeuGluArgGlnMetGlyTyrTyrLeuIleGln 275
DB 854 TATTCGGCTTATCGCTGGGAATTCAGTTTCGTGCGTTCGATGGCTACACTTATACA 913
QY 276 MetTyrIleProSerLeuLeuIleValIleLeuSerThrProValSerPheThrIleAsnMet 295
DB 914 ATCTCAATACCTCTGAGATGATCTGTATATATCATGGGTATCATTTTGGCTCAATCGC 973
QY 296 AspaAlaAlaProAlaArgValGlyLeuGlyIleThrThrValLeuThrMetThrGln 315
DB 974 AATGCAAGCCGCGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1033
QY 316 SerSerGlySerArgAlaSerLeuProLysValSerTyrValLysAlaIleAspIleThr 335
DB 1034 ATGTCGTCAACAATGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1093
QY 336 MetAlaValCysLeuLeuPheValPheAlaAlaLeuGluTyrAlaAlaIleAsnPhe 355
DB 1094 CTGGGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1153
QY 356 ValSerArgGlnHisLysGluThrLeuArgValGlnArgGlnArgGlnArgGln 375
DB 1154 ATGGCAAA-----CGAATTCAAATCGCAAAACAAGATTATATGCGC 1195
QY 376 GluGluAspIleIleGlnLysSerArgPheTyrPheArgGlyTyrGlyLeuGlyHisCys 395
DB 1196 ATCAAAAGATAGCGCAAGAAAAGCAACAGCTCGACGAGCGAACCAACAGCAGCGC 1255
QY 396 -----LeuGlnAlaArgAspGlyIleProMetGlyLysSerGlyIleTyrSerPro 412
DB 1256 AATCCCAATCCCAATGCAATGCGGCGGAGCC---GAGAGAGTGGCGCTTGACCCGCG 1312
QY 413 GlnProProAlaPro 417
DB 1313 GACCCCGAGAGACCC 1327
```

```
RESULT 14
US-08-072-064-5
Sequence 5, Application US/08072064
Patent No. 6008046
GENERAL INFORMATION:
APPLICANT: FRENCH-CONSTANT, RICHARD H.
APPLICANT: JACKSON, MEYER B.
TITLE OF INVENTION: DRUG AND PESTICIDE SCREENING
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: PETER G. CARROLL
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
```

```
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/072,064
FILING DATE: 19930602
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 770,881
FILING DATE: 04-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: CARROLL, PETER G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: OPHD-00574
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/705-8410
TELEFAX: 415/397-8338
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 2066 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: DNA
US-08-072-064-5
```

## Alignment Scores:

```
Pred. No.: 5 73e-82 Length: 2066
Score: 718.00 Matches: 146
Percent Similarity: 60.21% Conservative: 87
Best Local Similarity: 37.73% Mismatches: 118
Query Match: 31.88% Indels: 36
DB: 3 Gaps: 8
```

US-10-075-846-4 (1-431) x US-08-072-064-5 (1-2066)

```
QY 5 ValProAlaThrLeuSerPheLeuLeuThrThrLeuProGlyGlnValLeuLeuArg 24
DB 137 CTGCCCCGACACA---CCGCTGCTAACCATCTCG-----CTGCCCATCAAC 178
QY 25 ValAlaLeuAlaLysGluLys-----ValLysSerGlyThr 36
DB 179 ATGGCCCTGATTGACACAGCAAGCGCCACAAACGATCATACAGTGCACACGCGCAT 238
QY 37 LysGlySerGlnProMetSerProSerAspPheLeuAspLysLeuMetCylArgThrSer 56
DB 239 GCGGCGGACGATGCTGGGTGACGTAACATATCCGCTATTTCTGACTCTTTAGTGT 298
QY 57 GlyTyrAspAlaArgIleArgProAsnPhenLysGlyProProValAsnValThrCysAsn 76
DB 299 AGTTACGACAAAAGAGTAAGACCAATTAAGTGGTCCGCCCTGCGAGGTTGGCGTCA 358
QY 77 IlePheIleAsnSerPheSerSerValThrLysThrThrMetAspTyrArgValAsnVal 96
DB 359 ATGTATGTCCTCAGTATCAGTTCGGTTCCGAAAGTTCTAATGACTTTCACATTGGATT 418
QY 97 PheLeuArgGlnGlnIleProAsnAspProArgLysSerTyrArgGlyTyrProAsp---Asp 115
DB 419 TACTTTTCGTCATTTTGGACCGATCCCTTATAGCGTATAGAAAACGACCTGCTAGAA 478
QY 116 SerLeuAspLeuAspProSerMetLeuAspSerIleTyrLysProAspLeuPheAla 135
DB 479 ACAGTATCGCTGGATGATCAAGTTCATTAAGATATTTGGGTACCGACACCTTTTGTGA 538
QY 136 AsngluysgLyAlaAsnPhenHisGluValThrThrAspAsnLysLeuLeuArgIlePhe 155
DB 539 AATGAAAAAATCATATTTTCAATGTAACACAGTAATGAAATTCATACGTGAT 598
```





```
Db 598 ACGGTACGCGATGTGTCATGTAAGCCAAAGCCAAAGCCGCGAGTGGAG 651
OY 230 GlyLeuThrLeuProGlnPheIleu-----ArgaspGluLysAspLeu 244
Db 652 GATCGGAGCGCGCATTCACCATCTGGGTATGAGACCAATGACCGAAAGAG--- 708
OY 245 GlyCysCysThrLysHisLysThrGlyLysPheThrCysIleGluValLysPheHis 264
Db 709 -----CGGCTGGCCACTGGAGCTATACGCCCTCTGCTCATTTCAA 753
OY 265 LeuGluArgGlnMetGlyTyrTyrLeuIleGlnMetTyrIleProSerLeuIleVal 284
Db 754 CTGCACGGAATATCGGATCTTGTATTCCAAATTATCGCCAGCATTCGATCGTA 813
OY 285 IleLeuSerTrpValSerPheTrpIleAsnMetAspAlaProAlaArgValGlyLeu 304
Db 814 ATGCTGTCGTGGTCTGCTTCTGGATTAAACACAGCGAGAGTGGCGGTTGCATTG 873
OY 305 GlyIleThrThrValLeuThrMetThrThcGlnSerSerGlySerArgAlaSerLeuPro 324
Db 874 GGCATCACACGAGTGTCTCACCATGACCACTTAGCAGGGGTTCGCAGCTCACTGCCG 933
OY 325 LysValSerTyrValLysAlaIleAspIleTrpMetAlaValCysLeuLeuPheValPhe 344
Db 934 CGCATATCGTATGGAAGCGCATCGACATTATCTGTCATGTGCTTCTGTTGCTGTTTC 993
OY 345 AlaAlaLeuLeuGluTyrAlaAlaIleAsnPhe-----ValSerArg 358
Db 994 GCACCCCTCTGGAATACGCTGACGCTTAACATACTACTGAGGCCAAAGGGCTAA-AAA 1052
OY 359 GlnHisLysGluPheIleArgLeuArgArgGlnArgGlnArgGlnArgGluGluAsp 378
Db 1053 GAAATATAAGAAAGTCAAGAAATGTGTCCAGGCCAAGTCGGAAGAGTGAAGATCCGA 1112
OY 379 IleIleGlnGluSerArgPheTyrPheArgGlyTyrGly 391
Db 1113 GACGTGTTCAAGACAGACAGACATTATCGAGCTGCAGGA 1151
```

Search completed: July 1, 2003, 00:57:28  
Job time : 115.6 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: June 30, 2003, 23:07:05 ; Search time 229.481 Seconds  
(without alignments)  
2788.018 Million cell updates/sec

Title: US-10-075-846-4  
Perfect score: 2252  
Sequence: 1 MTLVPATISFLMLTLPGQ.....POPPADLRGCTETRLKYVD 431

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1055720 seqs, 742224136 residues  
Total number of hits satisfying chosen parameters: 2111440

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08  
Maximum Match 1008  
Listing first 45 summaries

Command line parameters:  
-MODEL-frame+ p2n.model -DEV-x1p  
-O/cg2\_1/USPTO.spool/US10075846/runat\_25062003\_163649\_5181/app.query.fasta.1.782  
-DB-published.Applications\_NA -OPMT-fastap -SUFFIX-p2n.rnpb -MINMATCH=0.1  
-LOOPL=0 -LOOEXT=0 -UNITS-bits -START=1 -END=1 -MATRIX-blosum62  
-TRANS-numa40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=100 -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MALEN=2000000000 -USER=US10075846.ecgn.1.1.102\_etunat\_25062003\_163649\_5181  
-NCU=6 -ICPU=3 -NO\_MAP -LARGQUERY -NEG\_SCORES=0 -WAIT -DSBLOCK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published\_Applications\_NA:\*

1: /cg2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cg2\_6/ptodata/2/pubpna/PCF\_NEW\_PUB.seq:\*  
3: /cg2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cg2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cg2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cg2\_6/ptodata/2/pubpna/PCFUS\_PUBCOMB.seq:\*  
7: /cg2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cg2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cg2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
10: /cg2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
11: /cg2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
12: /cg2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
13: /cg2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
14: /cg2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	2252	100.0	1640	US-10-075-846-3
2	2163	96.0	2565	US-10-075-846-1
3	1588	70.5	993	US-10-075-846-9
4	750	33.3	4621	US-09-969-844-13

5	749	33.3	1609	US-09-969-844-11	Sequence 11, Appl
6	742.5	33.0	1197	US-10-239-420-10	Sequence 10, Appl
7	728.5	32.3	1640	US-09-969-844-12	Sequence 12, Appl
8	726.5	32.3	1467	US-09-815-925-10	Sequence 10, Appl
9	719	31.9	3442	US-10-239-420-3	Sequence 3, Appl
10	719	31.9	3598	US-10-239-420-1	Sequence 1, Appl
11	698.5	31.0	1693	US-10-037-270-131	Sequence 131, App
12	694.5	30.8	1866	US-09-964-824-96	Sequence 96, Appl
13	680.5	30.2	1555	US-10-211-673-11	Sequence 11, Appl
14	670	29.8	1150	US-10-239-420-13	Sequence 13, Appl
15	655	29.1	2194	US-10-239-420-4	Sequence 4, Appl
16	652.5	29.0	1323	US-09-510-662A-38	Sequence 38, Appl
17	652.5	29.0	1323	US-09-778-320-38	Sequence 38, Appl
18	652.5	29.0	1323	US-09-910-689-38	Sequence 38, Appl
19	652.5	29.0	1323	US-10-010-742-38	Sequence 38, Appl
20	652.5	29.0	2593	US-10-106-698-2031	Sequence 2031, Ap
21	652.5	29.0	3282	US-09-954-531-1034	Sequence 1034, Ap
22	652.5	29.0	3282	US-09-825-301-75	Sequence 75, Appl
23	640.5	28.4	4077	US-10-239-420-6	Sequence 6, Appl
24	632.5	28.1	1192	US-09-808-483-5	Sequence 5, Appl
25	632.5	28.1	1269	US-09-808-483-3	Sequence 3, Appl
26	632.5	28.1	1272	US-09-808-483-7	Sequence 7, Appl
27	625	27.8	1997	US-10-037-270-256	Sequence 256, App
28	620	27.5	2819	US-09-818-657-1	Sequence 1, Appl
29	619.5	27.5	1707	US-10-211-673-7	Sequence 7, Appl
30	619	27.5	1987	US-10-037-270-331	Sequence 331, App
31	616	27.4	1398	US-09-893-321-1	Sequence 1, Appl
32	616	27.4	1568	US-09-893-321-5	Sequence 5, Appl
33	615	27.3	1458	US-09-808-483-11	Sequence 11, Appl
34	615	27.3	1608	US-09-808-483-9	Sequence 9, Appl
35	608	27.0	1404	US-09-765-069-1	Sequence 1, Appl
36	608	27.0	1422	US-09-742-311-1	Sequence 1, Appl
37	607	27.0	1263	US-09-765-069-7	Sequence 7, Appl
38	603	26.8	1179	US-09-765-069-3	Sequence 3, Appl
39	602	26.7	1038	US-09-765-069-9	Sequence 9, Appl
40	594.5	26.4	1444	US-09-839-446-5	Sequence 5, Appl
41	594.5	26.4	1444	US-09-898-570-5	Sequence 5, Appl
42	592.5	26.3	1438	US-09-839-446-7	Sequence 7, Appl
43	592.5	26.3	1438	US-09-898-570-7	Sequence 7, Appl
44	591	26.2	1417	US-09-839-446-3	Sequence 3, Appl
45	591	26.2	1417	US-09-898-570-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1  
US-10-075-846-3  
Sequence 3, Application US/10075846  
Publication No. US20030032608A1  
GENERAL INFORMATION:  
APPLICANT: Bristol-Myers Squibb Company  
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING A NOVEL GLYCINE RECEPTOR ALPHA SUBUNIT  
FILE REFERENCE: D0079 NP  
CURRENT APPLICATION NUMBER: US/10/075, 846  
PRIOR APPLICATION NUMBER: 2002-02-13  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 81  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 3  
LENGTH: 1640  
TYPE: DNA  
ORGANISM: homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(1293)  
US-10-075-846-3  
Alignment Scores:  
Pred. No.: 1.62e-278  
Score: 2252.00  
Percent Similarity: 100.00%  
Length: 1640  
Matches: 431  
Conservative: 0

Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

Oy		NethrThrlauValAProlaThlenseRPhelLeuleuTrpHleuProGlyIn	20
Dd	1	ATGCACACTCTTTTCCCTCGCAACCCTTCCTCTCTCTGTGGACCCCTGGCAGGGCAG	60
Oy		ValLeuleuArGvaAlaLeuAlaLysGluJuValLysSerGlyThrLysGlySerGln	40
Dd	61	GTCCTCCACAGGGTGCCCTTGGCAAATAAGAGAAGCAATCATCGAACCAAGGGGTCCAG	120
Oy		PrometSerProSerAspPheLeuAspLysIleuMetGlyAntgThSerGlyTyrrAspAla	60
Dd	121	CCCATTGCCCTCTGATTTCCTTGACAAACTTATGGGGCCGAACATCTGGATATGATCC	180
Oy		ArgIleArProAsnPhelLysGlyProProValAsnValThrCyAsnIllePheIleAsn	80
Dd	181	AGGATTCGGCCCAATTTAAAGGCCCACCCGAGAACGTGACCTGCACATCTTCATCAAC	240
Oy		SerPheSerSerValThrLysThrThrmetspyrArGvalAsnValPheLeuArgGln	100
Dd	241	AGTTTCACCTCGGTACCCACAACCAACAAATGACATCACGGGTGAATGTTCTTGGGGAA	300
Oy		GlnTrpAsnaSpProArGleuSerTyrrArgGluTyrrProAspAspSerLeuAspLeuasp	120
Dd	301	CAGTGAATGACCCACAGCCCTCTCCACGAGAAATTCCTGATGACTCTGTGACCTCGAT	360
Oy		ProserMetLeuaspSerIleTrpLysProAspleuPhePheAlaAsnGluLysGlyAla	140
Dd	361	CCCTTCATGCTGGACTCTATCTGTGGAGGACAGCCTCTTGTGGCTTAAGAGAAAGGGCC	420
Oy		AsnPheHisGluValThrThrAspaSnLysLeuLeuArgIlePheLysAsnGlyAsnVal	160
Dd	421	AAC TTCCTGAGGTGACACCGACGACAAACAATGACTGCGCCATCTTCACAAATGGGAATG	480
Oy		LeuTyrrSerIleArGleuThrLeuIleLeuSerCysIleuMetAspLeuLysAsnPhero	180
Dd	481	CTGTACACATCAGAGGTGACCCCTCATTTTGTCTGCTCGATGAGGACCTCAAGAACTTCCC	540
Oy		MetAspIleGlnThrCysThrMetGlndeuIuseSerSerIleLeuCysSerProLeu	200
Dd	541	ATGACATATCCAGAGGTGACGATGACGCTTGAGACTCATCATCTACTCTGCAGCCCTCG	600
Oy		ProserLeuSerLeuSerValGlyTyrrThrmetyAspLeuValPheGluTrpLeuGlu	220
Dd	601	CCAATCTGTCACTTTCAGTTGAGGTGACACCAATGAAGACCTCGTGTGGAGGCTGGAA	660
Oy		AspAlaProAlaValAlaGlnValAlaGluLysLeuThrLeuProGlnPheIleLeuArGras	240
Dd	661	GATGCTCTGCTGTCACAGTGGCGAGGGGCTGACTCGCCCAATTATCTTGGGGAT	720
Oy		GluIlysAspLeuGlyCysCysThrLysHisTyrrAsnThrGlyLysPheThrCysIleGlu	260
Dd	721	GAGAAAGGATCAGGCTGTGTACCAAGACATCAACACAGGAAATTCACCTGCATTCAG	780
Oy		ValLysPheHisLeuGluArgGlnMetGlyTyrrTyrrLeuIleGlnMetTyrrIleProSer	280
Dd	781	GTAAGAATTTCACCTGGAACGGCAGATGGGTACTATCATGATTCAAGATGTACATCCAGC	840
Oy		LeuIleuIleValIleLeuSerTyrrValSerPheTrpIleAsnMetAspAlaAlaProAla	300
Dd	841	CTACTCATCGATATCTGTCCTGGGTCTCTTGTGGATCAACATGGAAGCTGCCCCCTCC	900
Oy		ArgValGlyLeuGlyIleThrThrValLeuThrMetThrGlnSerSerGlySerArg	320
Dd	901	CGTGGGGCCCTGGGAGATCACCAACGCTGTCACCATGACACACCAAGCTGTGGCTCCGG	960
Oy		AlaSerLeuProLysValSerTyrrValLysAlaIleAspIleTrpMetAlaValCysLeu	340
Dd	961	GCCTCTTTGGCTTAAGGTGTCTCTAGCTGAAGCAATCGCATCTGGATGGCTGTGTGTGTG	1020

```

QY      341  leuphevalpheaiaaleuleuclntfyrhlaiaalleasnphelaserarglnhis 360
Db      1021  CTCYTTGGTGGCGCCCTTGCTGAGATGCTCCATAAATTTTGTTCGCGACCAT 1080
QY      361  LysgiupheileavleuAaArgaRgaRglaRfarglaRfgleuGluluaSP1le 380
Db      1081  AAAGATTCATACGACTTCGAAAGAGGAGGAGGCCAACGCTTGAGGAAGATATCATC 1140
QY      381  GlucluserarphetheyrPhearglyTyrglyleuGlyhisCysleuGlnAlaArgasp 400
Db      1141  CAAGAAAGCTCTTCTATTTCCGTGGCATGGCTTGCGCCACTGCTGCAGGCAAGAGAT 1200
QY      401  GllgylpPrometGluGlySerGlyilefyrSerProGlnpRopAlaapoluLeuArg 420
Db      1201  GGAGGTCCAAAGGAAGGTTCTGGGCATTTATAGTCCCAACCTCCAGGCCCTCTTCTAAG 1260
QY      421  GluGlyGluThrThrArgLysLeuTyxAlaSP 431
Db      1261  GAAGGAGAAACCGACGCGGAACCTCTACGTGGAC 1293

RESULT 2
US-10-075-846-1
; Sequence 1, Application US/10075846
; Publication No. US20030032608A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING A NOVEL GLYCINE RECEPTOR ALPHA SUBUNIT
; FILE REFERENCE: D0079 NP
; CURRENT APPLICATION NUMBER: US/10/075, 846
; CURRENT FILING DATE: 2002-02-13
; PRIOR APPLICATION NUMBER: US 60/269, 535
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 2565
; TYPE: DNA
; ORGANISM: homo sapiens.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1251)
US-10-075-846-1

Alignment Scores:
Pred. No.:          9.2e-267          Length:          2565
Percent Similarity: 2163.00          Matches:          416
Best Local Similarity: 96.52%        Conservative:      0
Query Match:        96.05%          Mismatches:      14
DB:                  9              Gaps:              1

US-10-075-846-4 (1-431) x US-10-075-846-1 (1-2565)
QY      1  MetfhrhrleuValProAlaThrleuSerPheuleuLeuTrPfrleuProGlyGln 20
Db      1  ATGACAACACTGTGTTCCTGCAACCCCTGCTTCCTCTCTCTGACCCCTGCCAGGGCAG 60
QY      21  ValuLeuLeuArgValAlaLeuAlaLysGluLValLysSerGlyThrLysGlySerGln 40
Db      61  GTCCCTCCCTCAAGGCTGGCTTGCAAAAGAGAGAAAGTCAAAATCTGGAACCAAGGCTCCAG 120
QY      41  PrometSerProSerAspPheleuAspLysIleuMetGlyArgThrSerGlyTyxAla 60
Db      121  CCCAGTCCCTCCCTCTGATTTCTTAGACAAACTTATGGGCGCAACATCTGCAATGAGAGCC 180
QY      61  ArgIleArpProAspNheLysGlyProProValAsnValThrCysAsnIlePheIleAsn 80
Db      181  AGGATGCGGCCCATTTTAAAGGCCCAACCGCTGACGACGACCTGCACATCTTATCATAC 240
QY      81  SerPheSerSerValThrLysThrThrMetAspTyxArgValaAsnValPheLeuArgGln 100

```

Db 241 AGTTTCAGCTCCGTACCAAGACCAATGAGTACCGGGTGAATGCTTTCTGGCGCA 300  
Qy 101 GlnTTPAsnAspProArgLeuSerTyrArgGluTyrProAspAspSerLeuAspLeu 120  
Db 301 CAGTGAATGACCCAGCCCTGCTCTACCGAGAAATCTCATGACTCTGAGACTTCGAT 360  
Qy 121 ProSerMetLeuAspSerIleTyrPlysProAspLeuPhePheAlaAsnGluTyrGlyAla 140  
Db 361 CCTCCATGCTGAGACTCTATCTGAAAGCCAGACCTCTCTTCTGCTAATGAAAGGGGCC 420  
Qy 141 AsnPheHisGluValThrThrAspAsnLysLeuLeuArgIlePheLysAsnGlyAsnVal 160  
Db 421 AACTTCATGAGGTGACACGAGCAACAACTTACTGGCATCTTCAAGATGGGAATGTG 480  
Qy 161 LeuTyrSerIleArgLeuThrLeuIleLeuSerCysLeuMetAspLeuLysAsnPhePro 180  
Db 481 CTGTACAGCTACAGGCTGAGCCCTCATTTTGTCTGCTGTATGGACCTCAAGAACTTCCC 540  
Qy 181 MetAspIleGlnThrCysThrMetGlnLeuGluSerSerIleLeuCysSerProLeu 200  
Db 541 ATGAGACTCCAGACCTGACGATGACAGCTTGAG----- 573  
Qy 201 ProSerLeuSerLeuSerValGlyTyrThrMetLysAspLeuValPheGluTyrPleuGlu 220  
Db 574 -----AGCTTTGGCTACACCATGAAAGACCTCGTGTGGAGTGGCTGGAA 618  
Qy 221 AspAlaProAlaValAlaGluGlyLeuThrLeuProGlnPheIleLeuArgAsp 240  
Db 619 GATGCTCTGCTGCTCCAGAGGCTGAGGGGCTGACTGCTGCCAGTTATCTTGGGGAT 678  
Qy 241 GlnLysAspLeuGlyCysThrLysHisTyrAsnThrGlyLysPheThrCysIleGlu 260  
Db 679 GAGAGAGATCTAGGCTGTTGTACCAAGCACTACACAGCAGGAATTCACCTGCATGAG 738  
Qy 261 ValLysPheHisLeuGluArgGlnMetGlyTyrTyrLeuIleGlnMetTyrIleProSer 280  
Db 739 GTAAGATTTCCACTGGAACGGCAGATGGCTACTATGATTCAGATGATGATCCCGAGC 798  
Qy 281 LeuLeuIleValIleLeuSerTyrPheIleAsnMetAspAlaIleProAla 300  
Db 799 CTACTCATCTCATCCCTGCTGCTGGGTCTCTTCTGATCAACATGATGCTGCCCTGCC 858  
Qy 301 ArgValGlyLeuGlyIleThrThrValLeuThrMetThrThrGlnSerSerGlySerArg 320  
Db 859 CGGTGGGGCTGGGACATCACCGCTGCTCACCATGACCAACAGAGCTGGGCTCCCGG 918  
Qy 321 AlSerLeuProLysValSerTyrValLysAlaIleAspIleTyrPheAlaValCysLeu 340  
Db 919 GCCTCTTTGGCTTAAGGTGCTCTACGTGAAGCAATCGCATCTGGATGCTGTCTCTG 978  
Qy 341 LeuPheValPheAlaIleLeuGluTyrValAlaIleAsnPheValSerArgGlnHis 360  
Db 979 CTCTTTGTCTGCTGCTGCTGCTGAGTATGCTGCCATAATTTTGTCTTCTGCTGACAT 1038  
Qy 361 LysGlnPheIleArgLeuArgArgArgGlnArgArgGlnArgLeuGluLysPheIle 380  
Db 1039 AAGAAATTCATACGACTTCGAAGAAGCAGAGGCCCAAGCTTGAGGAATATCATC 1098  
Qy 381 GlnGluSerArgPheTyrPheArgGlyTyrGlyLeuGlyHisCysLeuGlnAlaArgAsp 400  
Db 1099 CAAGAAAGTGGTTTCTAATTTCCGCGCTAATGCTTGGGCCATGCCGAGGCAAGGAT 1158  
Qy 401 GlyGlyProMetGluGlySerGlyIleTyrSerProGlnProProAlaProLeuLeuArg 420  
Db 1159 GGAAGTCAATGGAAGTCTGGCATTTATAGTCCCAACCTCCAGCCCTCTTCTAAGG 1218  
Qy 421 GlnGlyGluThrThrArgLysLeuTyrValAsp 431  
Db 1219 GAAGGAGAAACACGCGGAATCTACGTGAC 1251

RESULT 3

US-10-075-846-9

: Sequence 9, Application US/10075846

: Publication No. US20030032608A1  
: GENERAL INFORMATION:  
: APPLICANT: Bristol-Myers Squibb Company  
: TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING A NOVEL GLYCINE RECEPTOR ALPHA SUBUNIT  
: TITLE OF INVENTION: IN THE GASTROINTESTINAL TRACT, HORA4, and SPLICE VARIANT THER  
: FILE REFERENCE: D0079 NP  
: CURRENT APPLICATION NUMBER: US/10/075,846  
: CURRENT FILING DATE: 2002-02-13  
: PRIOR APPLICATION NUMBER: US 60/269,535  
: PRIOR FILING DATE: 2001-02-16  
: NUMBER OF SEQ ID NOS: 81  
: SOFTWARE: PatentIn version 3.0  
: SEQ ID NO 9  
: LENGTH: 993  
: TYPE: DNA  
: ORGANISM: homo sapiens  
: US-10-075-846-9  
  
Alignment Scores:  
Pred. No.: 1,29e-193 Length: 993  
Score: 1588.00 Matches: 309  
Percent Similarity: 89.86% Conservative: 1  
Best Local Similarity: 89.57% Mismatches: 1  
Query Match: 70.52% Indels: 34  
Gaps: 2  
DB:  
  
US-10-075-846-4 (1-431) x US-10-075-846-9 (1-993)  
Qy 1 MetThrThrLeuValProAlaThrLeuSerPheLeuLeuLeuThrPheLeuProGlyGln 20  
Db 1 ATGACAACTCTGTGCTCCGCAACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 60  
Qy 21 ValLeuLeuArgValAlaLeuAlaLysGluGluValLysSerGlyThrLysGlySerGln 40  
Db 61 GTCCCTCTCAGGGTGGCTTGGCAAAAGAGAGTCAAAATCTGAACCAAGAGGGTCCAG 120  
Qy 41 ProMetSerProSerAspPheLeuAspLysLeuMetGlyArgThrSerGlyTyrAspAla 60  
Db 121 CCCATGCCCCCTGATTTCTTAAGCAAACTTATGGGGCAACATCTGATATGATGCC 180  
Qy 61 ArgIleArgProAsnPheLysGlyProProValAsnValThrCysAsnIlePheIleAsn 80  
Db 181 AGGATTCGGCCCAATTTTAAAGGCCCAACCGTGACGTGACATGCAACATCTTCATCAAC 240  
Qy 81 SerPheSerValThrLysThrThrMet----- 90  
Db 241 AGTTTCAGCTCCATCCACCAAGACCAACATGGCTTGGGCCCCCTGGGAATGGCAATGTT 300  
Qy 91 -----AspTyrArgValAsnValPheLeuArgGln 100  
Db 301 TCTGAAGGCCCATATCTGCAACCTCCAGAGACTACCGGGTGAATGCTTCTTGGCGCA 360  
Qy 101 GlnTTPAsnAspProArgLeuSerTyrArgGluTyrProAspAspSerLeuAspLeu 120  
Db 361 CAGTGAATGACCCAGCCCTGCTCTACCGAGAAATCTCATGACTCTTCTGAGACTTCGAT 420  
Qy 121 ProSerMetLeuAspSerIleTyrPlysProAspLeuPhePheAlaAsnGluTyrGlyAla 140  
Db 421 CCTCCATGCTGAGACTCTATCTGAAAGCCAGACCTCTCTTCTGCTAATGAAAGGGGCC 480  
Qy 141 AsnPheHisGluValThrThrAspAsnLysLeuLeuArgIlePheLysAsnGlyAsnVal 160  
Db 481 AACTTCATGAGGTGACACGAGCAACAACTTACTGGCATCTTCAAGATGGGAATGTG 540  
Qy 161 LeuTyrSerIleArgLeuThrLeuIleLeuSerCysLeuMetAspLeuLysAsnPhePro 180  
Db 541 CTGTACAGCTACAGGCTGAGCCCTCATTTTGTCTGCTGTATGGACCTCAAGAACTTCCC 600  
Qy 181 MetAspIleGlnThrCysThrMetGlnLeuGluSerSerIleLeuCysSerProLeu 200  
Db 601 ATGAGACTCCAGACCTGACGATGACACTTGAG----- 633  
Qy 201 ProSerLeuSerLeuSerValGlyTyrThrMetLysAspLeuValPheGluTyrPleuGlu 220







Db	121	AGAAATTGATGACGCTCACCTGATGAGTACACAGTCGCAAAAGAGCTTCACAGAGACAGTGGCG	180
Oy	104	ASPRQAIrLeuSerTyrArgLyuTyrProAspSerLeuAspLeuApproSerMet	123
Db	181	GACGAGAGATCTCAGTACGACGACTTGGGGCGCCAGGTTCCCTACCTGACGCTCCAGCA	240
Oy	124	LeuAspSerTLeuTyrPyrProAspPhePheAlaAsnGluTyrSgIyAlaAsnPheHis	143
Db	241	CCGAGACAGCTTGGAAAGCCGGAGCCTGTTTTCTCCACGAGAAAGAGGACACTTCAC	300
Oy	144	GluValThrTyrAspAsnLysLeuLeuArgTLeuPheLysAsnGlyAsnValLeuTyrSer	163
Db	301	AACATCATCATGCCCAACAGCTGCTTCAGCAATCATCCCAACGGCGAGCTTCCTTACG	360
Oy	164	ILeArgLeuThrLeuLeuLeuSerCysLeuLeuTAspLeuTAspPheProMetAspIle	183
Db	361	ATGAGAAATATCTTGSTGGCTTTCATGTCGATGAACCTGAAATTTTATCTTGGATAA	420
Oy	184	GlnThrCysThrMetGlnLeuGlnLeuSerSerIleLeuCysSerProLeuProSerLeu	203
Db	421	CAAAATGCTCTATCT-----GTC	438
Oy	204	SerLeuSerValGlyTyrThrMetLysAspLeuValPheGluTyrPheGluAspAlaPro	223
Db	439	ATGGTAGGATATGGATGATACAAACAGACACCTGGGTTCATGATGAAGAGCGGATCCT	498
Oy	224	AlaValGlnValAlaGlnLysLeuThrLeuProGlnPheIleLeu---ArgAspGluTyr	242
Db	499	---GTACAGGTGCACAAAATATCTCCACATTCGCCAGCTTCACGCTGGAAGAGTTTCAAC	555
Oy	243	AspLeuGlyCysCysThrLysHisTyrAsnThrGlyLysPheMetCysIleGluValLys	262
Db	556	GACTAC-----TGCAACGAGTGGGACCAACATGGCGATGACACTGCTTGGCGGTGAC	609
Oy	263	PheHisLeuGlnArgLysMetGlyTyrTyrLeuIleGlnMetCyrIleProSerLeu	282
Db	610	CTGGTGTCAAGCGGAGTTCAGCTACATCACTGATCCAGATCTACATCCGCTGTCATG	669
Oy	283	IleValIleLeuSerTyrPAlaSerPheThrPileAsnMetAspAlaAlaProAlaArgVal	302
Db	670	CTGGTCATGTCGTGCTGGGTGTCGTTCGGCTGCGACCCACCTGATGCCGGCGCGAGTG	729
Oy	303	GlyLeuGlyIleThrThrAlaLeuThrMetThrThrGlnSerSerGlySerArgAlaSer	322
Db	730	TCGCTGGGCTCACCCACCTGCTCACCATGGCCACGCAAGATATGGGATCAACGCTCG	789
Oy	323	LeuProLysValSerTyrValLysAlaIleAspIleThrPheTAlaValCysLeuLeuPhe	342
Db	790	CTGCTCCGCTGTTCTTACACCAAGGCAATGACGTGGACGGCGTGTCTGTCACCTTC	849
Oy	343	ValPheAlaAlaLeuLeuGlnTyrAlaAlaIleAsnPheValSerArgGlnHisLysGlu	362
Db	850	GTAATGGCGCGCTCTCTGAGTTGGCCCTGGTGCAATCAGCGCTCGCATTCACGC	909
Oy	363	PheIleArgLeuArgTArgArgLysIleArgGln-----ArgLeuGlnGlu	377
Db	910	CGGACAGACATGACAGAGACAGAAAGAGAAATGGAGCTGACCGCCCTTGAGATCG	969
Oy	378	AspIleIleGlnGlnLysSerArgPheTyrPheArgGlyTyrGlyLeuGlnHis	394
Db	970	GACACACCTGGAGGACGCGCCACCAACGTTCCGATGAGCGCGCTGGTGCAC	1020
RESULT 7			
US-09-969-844-12			
; Sequence 12, Application US/09969844			
; Publication No. US20020192776A1			
; GENERAL INFORMATION:			
; APPLICANT: Xiao-Zhou Michelle Wang			
; APPLICANT: Xavier Georges Sarda			
; APPLICANT: Michael David Tomalski			
; APPLICANT: Vincent Paul Mary Wingate			
; TITLE OR INVENTION: Heliotrans Glutamate Receptor			
; FILE REFERENCE: A32815-1 072667 .0178			

```

; CURRENT APPLICATION NUMBER: US-09/969,844
; CURRENT FILING DATE: 2001-10-03
; NUMBER OF SEQ. ID NOS: 14
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 1640
; TYPE: DNA
; ORGANISM: Artificial Sequence
FEATURE:
; OTHER INFORMATION: PCR amplified fragment of Helicobis DNA cloned
US-09-969-844-12

Alignment Scores:
Pred. No.:          7,06e-83           Length:      1640
Score:             728.50            Matches:       156
Percent Similarity: 61.108           Conservative:  56
Best Local Similarity: 44.96%        Mismatches:   110
Query Match:       32.35%            Indels:       25
DB:                Gaps:              8

US-10-075-846-4 (1-431) x US-09-969-844-12 (1-1640)

Oy    31 GluVallysserGlyThrlysglyserGlnProMetSerProSeraspPheLeuAspLys 50
Db    154 GAATGCATGACCGTGGAGCATCCACTTTCCAGAAGAGAGAACAGATCCTGGATCAG 213
Oy    51 LeuMeGclYarGthrserGlyTyraSpAlaArgIleatPro-----AsnPhelys 67
Db    214 ATCTTGACC---CCCCGAGAGTAACGACGCCAGATCAACCTCGGGATCACGGCACT 270
Oy    68 Gly--ProboVAlasnaValThrcysasnIlePheIleasnSerPheSerValThr 86
Db    271 GGCTATGCCGCAACGTTAGTCATGTCACAATGATTCTACGGCTCATCAGCAAAATGAT 330
Oy    87 LysThrThreIsptyrArValasnaValPheLeuAgsInghrTPAsnAspProArg 106
Db    331 GATTACAAAATGGGAATACCTCCGTACAAATTAACGTTCCGGGAACATGGATTGATGAACGG 390
Oy    107 LeuSerTyriArgIuTrProAspAspSerLeuAspLeuAspProSerMetLeuAspSer 126
Db    391 CTCAAATTCAATATCTTGAGAGTCGCCCAATATCCTGCACATGCATGACGTAAGCCAACGA 450
Oy    127 IleTrplysrProAspleuPhePheAlasngIulysgIylaaSnPheHisgluValThr 146
Db    451 GTCTGATGCTGATATCTATTCTTCCCAACGAGGAAGAGTCAATTTCACAACATCATC 510
Oy    147 ThrAspAnlysLeuLeuArGtllePheIlysaSngIyaSnvalLeuTySerIleArgLeu 166
Db    511 ATGCCCAACCTGTACATCCGATCTTCCCACAAGCGCAACGTGCTACAGATCCGAATC 570
Oy    167 ThrLeuIleLeuSerCysLeuMetLaspleuLyasaNphePrometAspIlegInThrCys 186
Db    571 TCCCTGACGCTGCTCGGCCCACTGAACCTCAAGTTGTCGCCCTGTAAAGCAACCTGC 630
Oy    187 ThMeTgIleuLeuIubserSerIleLeuCysserProIeaProserLeuSerLeuSer 206
Db    631 TCCTGTCGGCTG-----GCTAGT 648
Oy    207 ValGIlyTrHemetylaspleuValPhegluTrPleugluAspAlaProAlaValGln 226
Db    649 TATGTTGGACCCACACAGACAGATTAGTGTTCATGGAAGAAAGCGACCCG--GTCCAG 705
Oy    227 ValAlagluGlyLeuThrLeuProGlnPheIleLeuArGsaRglulysAspLeu--Gly 245
Db    706 GTGTGAAAAAATTTAACCTGCTCGGTTCACGCG---GAGAAGTTCCTCACTGAC 759
Oy    246 CysCysThrLysHisTyraSnhrGlylsPheThrCysIlegluValLysPheHisLeu 265
Db    760 TACTGCAACAGTAAGACTAATACCGGTGAATACATTCGCTGAAGGTAGACCTGCTTTC 819
Oy    266 GluArgIleMetGlyTyrrTyrlleuIleGlnMetTyrlleProSerLeuIleValIle 285

```

Db 820 AACGGAGTTCAGTACTACCTGATCCAGATCTACATTCCTGCTGCATGCTGTCATC 879  
Oy 286 LeuSerTrpValSerPheTrpIleAsnMetAspAlaIleProIleArgValIleGlyLeu 305  
Db 880 GTGCTGGGTGCTCTGCTGCTGACACGAGGAGCTGTGCGAGGGCTCTCAGTACGA 939  
Oy 306 IleThrThrValLeuThrMetThrThrGlnSerSerGlySerArgAlaSerLeuProIys 325  
Db 940 GTACAGCCTTACTTACATGAGGGACCACTGCTCAGGCGATCAACGCGTCCCTACCC 999  
Oy 326 ValSerTrpValIleAlaIleAspIleTrpMetAlaValIleCysLeuLeuPheVal 345  
Db 1000 GTGCTCTACACGAAACGATGATGCTGACTGCTGATCTCATTCATTCGATTCGGA 1059  
Oy 346 AlaLeuLeuGlnIleValIleAlaIleAsnPheValSerArg-----GlnHisIysGlnPhe 363  
Db 1060 GCCCTCTACAGTGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1119  
Oy 364 IleArgLeuArgArgArgGln 370  
Db 1120 ATGAGAAACGAGACGCGAG 1140

## RESULT 8

US-09-815-925-10  
: Sequence 10, Application US/09815925  
: Patent No. US20020127199A1  
: GENERAL INFORMATION:  
: APPLICANT: Tang, Y. Tom  
: APPLICANT: Zhou, Ping  
: APPLICANT: Goodrich, Ryle  
: APPLICANT: Asundi, Vinod  
: APPLICANT: Yang, Yonhong  
: APPLICANT: Zhang, Jie  
: APPLICANT: Mehrman, Tom  
: APPLICANT: Dimauc, Radoje T.  
: TITLE OF INVENTION: No. US20020127199A1el Nucleic Acids and  
: FILE REFERENCE: 787CIP2H  
: CURRENT APPLICATION NUMBER: US/09/815,925  
: PRIOR FILING DATE: 2001-03-22  
: PRIOR APPLICATION NUMBER: 09/560,875  
: PRIOR FILING DATE: 2000-04-27  
: PRIOR APPLICATION NUMBER: 09/496,914  
: PRIOR FILING DATE: 2000-02-03  
: NUMBER OF SEQ ID NOS: 11  
: SOFTWARE: PL-FL-Genes Version 2.0  
: SEQ ID NO 10  
: LENGTH: 1467  
: TYPE: DNA  
: ORGANISM: Homo sapiens  
: FEATURE:  
: NAME/KEY: CDS  
: LOCATION: (46)..(1464)  
US-09-815-925-10

## Alignment Scores:

Pred. No.: 1,06e-82 Length: 1467  
Score: 726.50 Matches: 157  
Percent Similarity: 58.59% Conservative: 68  
Best Local Similarity: 40.89% Mismatches: 111  
Query Match: 32,26% Indels: 48  
DB: 10 Gaps: 7

US-10-075-846-4 (1-431) x US-09-815-925-10 (1-1467)

Oy 4 LeuValProIleAlaThrLeuSerPheLeuLeuLeuTrpThrLeuProGlnValLeuLeu 23  
Db 58 CTCCTGGAGCTCTCTGCTGCCATCTGCTCTCTGACCTG----- 99  
Oy 24 ArgValAlaLeuAlaIleValIleGlnValIleSerGlyThrIysGlnIleProMetSer 43  
Db 100 -----GGACCGAGGCTCTGAGCCCGAGT 126

Oy 44 ProSerAsp-----PheIleuAspIleuMetGlyArg 54  
Db 127 GTACAGCATCCCGGAACATGCTCTTGTGACGAGACGCTGACAACTGTTG----- 180  
Oy 55 ThrSerGlyTrpAspAlaArgIleArgProAsnPheIysGlyProProValAsnValThr 74  
Db 181 ---AAGGCTACAGACATTCGCTTAAGACCCGACTTGGGGGTCGCCGCTCGTGGGG 237  
Oy 75 CysAsnIlePheIleAsnSerPheSerSerValIleThrThrIleAspIleArgVal 94  
Db 238 ATGACATATGACATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 297  
Oy 95 AsnValPheLeuArgGlnIleTrpAsnProArgLeuSerTrpArgIleProAsp 114  
Db 298 ACCATGATTTTAAACAAATTTTGAAGATAAAGGCTGCCATTTCTGGGATCCCT--- 354  
Oy 115 AspSerLeuAspLeuAspProSerMetLeuAspSerIleTrpIleProAspLeuPhe 134  
Db 355 CTCACCTCACCGCTGACATTCGAGTGGTACACCTATGAGTGGTCCGACACATATTTTC 414  
Oy 135 AlaAsnGlnIysGlyIleAsnPheHisGlnValIleThrAspAsnIleLeuArgIle 154  
Db 415 TTAATGACAAAGATCATTTTGTGATGAGTACAGTGAAGAAACGATGATGCTGCTT 474  
Oy 155 PheIleAsnGlyAsnValLeuTrpSerIleArgLeuThrLeuIleLeuSerCysLeuMet 174  
Db 475 CACCTGATGAGGACATGCTGCTGATGGCTGAGATACACCGACGACACGATCATGATG 534  
Oy 175 AspLeuIleAsnIlePheProMetLeuAspIleGlnThrCysThrMetGlnLeuGlnSer 194  
Db 535 GACCTCAGAGATACCCCTGAGCAGACGAACTGACTGGAATTTGAA----- 585  
Oy 195 IleLeuCysSerProLeuProSerLeuSerLeuSerValGlyIleThrMetIleAspLeu 214  
Db 586 -----AGCTATGCTGCTACACACGATGACATTT 612  
Oy 215 ValPheGlnTrpLeuGlnAspAlaProIleValIleGlnValIleGlnValLeuTrpLeuPro 234  
Db 613 GACTTTTACTGGGAGCGGAGCGGACAAAGCTGTTACCGAGTGAAGAGATGAGTCCCG 672  
Oy 235 GlnPheIleLeuArgAspIleGlnIleCysThrIleAsnIleArgIle 254  
Db 673 CACTTCTCCATCGTGGAGCACCT---CTGCTCGAGAAATGTTGCTTCCGACAGGT 729  
Oy 255 LysPheThrCysIleGlnValIlePheHisLeuGlnArgGlnMetGlyTrpLeuIle 274  
Db 730 GCCTATCTCTGACTGCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 789  
Oy 275 GlnMetTrpIleProSerLeuLeuIleValIleLeuSerTrpValSerPheTrpIleAsn 294  
Db 790 CAGACTTATATGCTCTCTATCTGATGATGATTTCTGCTGCTGCTGCTGCTGCTGCTG 849  
Oy 295 MetAspAlaIleProIleArgValIleGlyIleThrThrValLeuThrMetThrThr 314  
Db 850 TATGATGATGATCTCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 909  
Oy 315 GlnSerSerGlySerArgAlaSerLeuProIleValSerTrpValIleAlaIleAspIle 334  
Db 910 ATCAACACCCACCTTGGGAGCTTGGCCCAAAATGCCCATATCTCAAAAGCATATGACATG 969  
Oy 335 TrpMetAlaValIleCysLeuLeuPheValIleAlaIleLeuGlnIleValIleAsn 354  
Db 970 TACCTTATGGGCTGCTTCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1029  
Oy 355 PheVal-SerArgGlnHis-----LysGlnPheIleArgLeuArgArgGlnArgArg 372  
Db 1030 TACATTTTCTTGTGAGAGAGCGCTCAAAAGCAGAAAGAGCTTGCAGAAAGACAGCCAA 1089  
Oy 372 GlnArgLeu 375  
Db 1090 GCAAGAAATG 1099

## RESULT 9

```

US-10-239-420-3
: Sequence 3, Application US/10239420
: Publication No. US20030096984A1
: GENERAL INFORMATION:
: APPLICANT: Cully, Doris F.
: APPLICANT: Zheng, Yinqong
: TITLE OF INVENTION: DNA MOLECULES ENCODING LIGAND GATED ION
: FILE REFERENCE: 20629P
: CURRENT APPLICATION NUMBER: US/10/239,420
: PRIORITY FILING DATE: 2002-09-23
: PRIOR APPLICATION NUMBER: PCT/US01/09956
: PRIOR FILING DATE: 2001-03-28
: PRIOR APPLICATION NUMBER: 60/193,935
: PRIOR FILING DATE: 2000-03-31
: NUMBER OF SEQ ID NOS: 13
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 3
: LENGTH: 3442
: TYPE: DNA
: ORGANISM: Dermacentor variabilis
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (32)...(1225)
: US-10-239-420-3

Alignment Scores:
Pred. No.: 4.08e-81 Length: 3442
Score: 719.00 Matches: 150
Percent Similarity: 62.01% Conservative: 72
Best Local Similarity: 41.90% Mismatches: 98
Query Match: 31.93% Indels: 38
DB: 9 Gaps: 10

US-10-075-846-4 (1-431) x US-10-239-420-3 (1-3442)

OY 9 LeuSerPheLeuLeuLeuTrp-----ThrLeuProGlyGlnValLeuLeuArg 24
Db 38 CTTTCAGCGCTGCAACGCTGCGCGCTTGCCTCAGCTTGCTCC-----CTCCTCAGG 88
OY 25 ValAlaLeuAlaValysGlnGluVallySerGlyThrlySerGlnProMetSerPro 44
Db 89 ACGACGCTCGCGCAGGAAGCGGCTCAACGAGCG----- 124
OY 45 SerAspPheLeuAspLysLeuMetClyArgThrSerGlyTyraSpAlaArgIleArgPro 64
Db 125 CTGATGACGCTGGAAGACCTTGACGACTTATTAGAACCTTGAACCGCGCGCTTCC 184
OY 65 ---AsnPhelysGlyProProValAsnValThrCysAsnIlePheIleAsnSerPheSer 83
Db 185 ACGACACACTTGGAACCGCAACAAAGCTGGCTTGCAGAACTCTACATCGCAGCTTCGG 244
OY 84 SerValThrLysThrMetAspTyrArgValAsnValPheLeuArgGlnIleTrpAsn 103
Db 245 TCCTAATAATCCAGCCACACATGAGCTATGAGTTGATCTTTATTTCGGCAGACTTGCGCA 304
OY 104 AspProArgLeuSerTyrArgGluTyrProAspAspSer-----LeuAspLeu---Asp 120
Db 305 GATGATCGCTTGAGC-----AGCCCAACGATATCAGGCCCTCGACCTCAATGAT 355
OY 121 ProSerMetLeuAspSerIleTrpLysProAspLeuPhePheAlaAsnGlnLysGlyAla 140
Db 356 CCAAGAGCTGTCAGCGCTATATGGAACAGAGATATCTTCGCAAAATGCAAAACACGCA 415
OY 141 AsnPhenIsgValThrThrAspAsnLysLeuArgIlePheLysAsnGlyAsnVal 160
Db 416 GAGTTCCAAATATGTCACAGTACCTTAATGTTACTGGTCCGCTTAACCGAAGGAAGATT 475
OY 161 LeuTyrSerIleArgLeuThrLeuIleLeuSerCysLeuMetAspLeuLysAsnPhePro 180
Db 476 CTATTCATGCTCAGGCTCAAGCTTAAGTTTGCATGATGATGATGATGATGATGATGAT 535
OY 181 MetAspIleGlnThrCysThrMetGlnLeuGlnSerSerIleLeuCysSerProLeu 200

```

```

Db 536 ATGACATCCCAAGTTTGACGATCGAACTCCCTCATTTCTCG----- 577
OY 201 ProSerLeuSerLeuSerValGlyTyrThrMetLysAspLeuValPheGlnTrpLeuGlu 220
Db 578 -----AAAAACAACCAAGAACTGCAATCGAGTGGTCCGAT 613
OY 221 AspAlaProAlaValGlnValAlaGluGlyLeuThrLeuProGlnPheIleLeuArgAsp 240
Db 614 ACCAATCCGATTAATA---CTTTGGAAGGCTGGAAGTTACACAGTTCCGAGATTGACAAAT 670
OY 241 GluLysAspLeuGlyCysCysThrLysHisTyrAsnThrGlyLysPheThrCysIleGlu 260
Db 671 ---ATAAATACGTCAATCTGCATGAGAGAAATTTTCACATCGCAGAGTACAGCTGCTGAG 727
OY 261 ValLysPheHisLeuGluArgGlnMetGlyTyrTyrIleLeuIleGlnMetCysIleProSer 280
Db 728 GCCGACTTCCACTTGCAAGCGGTCACTGGCTACACATGGTGCTGATCTGCTTACCA 787
OY 281 LeuLeuIleValIleLeuSerTrpValSerPheTrpIleAsnMetAspAlaAlaProAla 300
Db 788 GTGCTCATCGTGCATCTCTGAGGTGCTTCTGCTGCTGCAAGTTCATTCATTCGCGG 847
OY 301 ArgValGlyLeuGlyIleThrThrValLeuThrMetThrThrGlnSerGlySerArg 320
Db 848 CGCACACACTGGCGGCTGACAGCGCTGCTCACTATTCTTCCAGAGCGCTCCGATATACAG 907
OY 321 AlaSerLeuProLysValSerTyrValLysAlaIleAspIleTrpMetAlaValCysLeu 340
Db 908 TCCACCTTGCTCCGCTCTCACTTACCTGAAGCAATTCATGCTGAGAGGAGCGCTGACG 967
OY 341 LeuPheValPheAlaAlaLeuLeuGluTyrAlaAlaIleAsnPheValSerArg 358
Db 968 GCCTTCGCTCTGCTGCGACCTACTGAGTTCACCGCTGCTGACGCTCCGCGGCGAG 1021

```

```

RESULT 10
US-10-239-420-1
: Sequence 1, Application US/10239420
: Publication No. US20030096984A1
: GENERAL INFORMATION:
: APPLICANT: Cully, Doris F.
: APPLICANT: Zheng, Yinqong
: TITLE OF INVENTION: DNA MOLECULES ENCODING LIGAND GATED ION
: FILE REFERENCE: 20629P
: CURRENT APPLICATION NUMBER: US/10/239,420
: PRIORITY FILING DATE: 2002-09-23
: PRIOR APPLICATION NUMBER: PCT/US01/09956
: PRIOR FILING DATE: 2001-03-28
: PRIOR APPLICATION NUMBER: 60/193,935
: NUMBER OF SEQ ID NOS: 13
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1
: LENGTH: 3598
: TYPE: DNA
: ORGANISM: Dermacentor variabilis
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (170)...(1363)
: US-10-239-420-1

Alignment Scores:
Pred. No.: 4.39e-81 Length: 3598
Score: 719.00 Matches: 150
Percent Similarity: 62.01% Conservative: 72
Best Local Similarity: 41.90% Mismatches: 98
Query Match: 31.93% Indels: 38
DB: 9 Gaps: 10

US-10-075-846-4 (1-431) x US-10-239-420-1 (1-3598)

OY 9 LeuSerPheLeuLeuLeuTrp-----ThrLeuProGlyGlnValLeuLeuArg 24

```

```

Db      176 CTTTACGCCCTGAACGTGGCGCGCTTCGCTGCACCTTCTCC-----CTCCCTCAGG 226
Oy      25 ValAlaLeuAlaLysGlnGluValLysSerGlyThrLysGlySerGlnPrometSerPro 44
Db      227 ACGAGCCCTGGCCGAGAAAGCGGCTCAAAACGGAGCG-----262
Oy      45 SerAspPheLeuAspLysLeuMetGlyArgThrSerGlyThrAsnAlaArgLeuArgPro 64
Db      263 CTGGATGACCTGAGAGCTTGCAGCTTATTAAGAACCTAGACCGGCTGCCCTTCC 322
Oy      65 ---AsnPhelYsgLysProProValAsnValThrCysAsnLlePheLleAsnSerPheSer 83
Db      323 ACGACACACTGGGAGACGCCAACAAAGTGGCTTGGCAATCTACATACGACCTTCGGG 382
Oy      84 SerValThrLysThrThrAspTyrArgValAsnValPheLeuArgGlnGlnTyrPasn 103
Db      383 TCCATTAATCCAGCCCAATGGACTATGAGGTTGATCTTATTGGCGGACACTTGGCAA 442
Oy      104 AspProArgLeuSerTyrArgGluTyrProAspAspSer-----LeuAspLeu---Asp 120
Db      443 GATGATGGCTTACG-----ACCCCAACGATTCAGGCGCCCTGGACCTCATGAT 493
Oy      121 ProSerMetLeuAspSerLleTyrLysProAspLeuPhePheAlaAsnGluLysGlyAla 140
Db      494 CCAAACCTGGTGCAGCGTATATGAAACCGGAAGTATCTTCCCAATGCCAAACACGCA 553
Oy      141 AsnPhelHisGluValThrThrAspAsnLysLeuLeuArgLlePheLysAsnGlyAsnVal 160
Db      554 GAGCTTCAATATGCTACACTACCTAATGTACTGCTGCTTACCGCTTAAACCGAAGAAAGT 613
Oy      161 LeuTyrSerLleArgLeuThrLleuLleLeuSerCysLeuMetAspLeuLysAsnPhaPro 180
Db      614 CTATACATGCTCGTCAAGCTCAAGTTTGCATGTATGATGATGATTTATTCCTTCC 673
Oy      181 MetAspLleGlnThrCysThrMetGlnLeuGlnSerSerLleLeuLysSerProLeu 200
Db      674 ATGGACTCCCAAGTATGACACATGCAACGACCTGCTTCCTCC-----715
Oy      201 ProSerLeuSerLeuSerValGlyTyrThrMetLysAspLeuValPheGlnTyrPleuGlu 220
Db      716 -----AAACACACCGAAGAACTGCATCTGGAGTGGTCTGAT 751
Oy      221 AspAlaProAlaValGlnValAlaGlnGlyLeuThrLeuProGlnPheLleLeuArgAsp 240
Db      752 ACCAATCCGATATA---CTATTCGAAGGCGCTGAGTTACCAATTCGAGATTCAGAAT 808
Oy      241 GlnLysAspLeuGlyCysCysThrLysHisTyrAsnThrGlyLysPheThrCysLleGlu 260
Db      809 ---ATAAATACGTCAATCTGCATGAGAAATTTCAATCGAGACATACACTCCCTGAG 865
Oy      261 ValLysPheHisLeuGlnArgGlnMetGlyTyrTyrLeuLleGlnMetTyrLleProSer 280
Db      866 GCGGACTTCCACTTCGACGCGGTACCTGGCTACCAATGTCGATCGATTCGCTTACA 925
Oy      281 LeuLeuLleValLleLeuSerTyrValSerPheTyrLleAsnMetAspAlaAlaProAla 300
Db      926 GTCTCATGCTGTCATCTGCTGCTGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 985
Oy      301 ArgValGlyLeuGlyLleThrThrValLeuThrMetThrThrLysSerGlySerArg 320
Db      986 CGGACACACACTGGCGCTGACGAGCTGCTCATTCTTCCTCAAGGCTCCGCTATACAG 1045
Oy      321 AlaSerLeuProLysValSerTyrValLysAlaLleAspLleThrPheAlaValCysLeu 340
Db      1046 TCCAACTTGGCTCGGCTCATACGTGAAGCAATGATGTGTGATGGAGGCTCTGCACG 1105
Oy      341 LeuPheValPheAlaAlaLeuLeuGluTyrValAlaAlaLleAsnPhaValSerArg 358
Db      1106 GCCTTGTGTTCTCGGACACTAGGATTCACCGCTGCTGCTGCTGCTGCTGCTGCTG 1159

```

```

; Sequence 131, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Li, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Tang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yundun
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Tillinghast, John
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/10/037, 270
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552, 317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488, 725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pL_FL_genes Version 1.0
; SEQ ID NO 131
; LENGTH: 1693
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (88)..(1512)
; US-10-037-270-131

Alignment Scores:
Pred. No.: 5,3e-79 Length: 1693
Score: 698.50 Matches: 145
Percent Similarity: 60.34% Conservative: 68
Best Local Similarity: 41.08% Mismatches: 115
Query Match: 31.02% Indels: 25
Gaps: 6

US-10-075-846-4 (1-431) x US-10-037-270-131 (1-1693)
Oy      10 SerPheLeuLeuLeuThrPheLleuProGlyGlnVal-----LeuLeuArgValAla 26
Db      76 TCGAGATTACTAATGTCGACAGTCAAAATCGAGAGTCTGGGCTTCTCTTTCCT 135
Oy      27 LeuAlaLysGlnGluValLysSerGlyThrLysGlySerGlnPrometSerProSer--- 45
Db      136 GTGATGATTACCATGCTGCTTCTGCACACAGACCAATGAACCAACCAACATGTCATAG 195
Oy      46 -----AspPheLeuAspLysLeuMetGlyArgThrSerGlyTyrAsnAlaArgLeuArg 63
Db      196 GTGAAGACGACATGACAGATTGCTC-----AAAGGATATGACATTCTGCTTGGCG 246
Oy      64 ProAsnPhelYsgLysProProValAsnValThrCysAsnLlePheLleAsnSerPheSer 83
Db      247 CCGGACTTCCGAGGCGCCCGCTTCGACGCTTGGATGCGATGATGATGCTCCGACATGAC 306
Oy      84 SerValThrLysThrThrMetAspTyrArgValAsnValPheLeuArgGlnGlnTyrPasn 103
Db      307 ATGCTCTCCGAAGTCAATATGATATACACTCACCATGATTTCCACGACGCTTGGAAA 366
Oy      104 AspProArgLeuSerTyrArgGluTyrProAspAspSerLeuAspLeuAspProSerMet 123
Db      367 GACAAAGGCTTCTTATTTCTGGAAATCCCA---CTGAACCTCACCTAGACATATAGGTA 423

```

[illegible]

Pred. No.:	2, 04e-78	Length:	1866
Score:	694.50	Matches:	144
Percent Similarity:	60.06%	Conservative:	68
Best Local Similarity:	40.79%	Mismatches:	116
Query Match:	30.84%	Indels:	25
DB:	10	Gaps:	6

  

US-10-075-846-4 (1-431) x US-09-964-824A-96 (1-1866)	
QY	10 SerPheLeuLeuLeuThrPheLeuProGluGlnVal-----LeuLeuArgValAla 26
Db	20 TCAGAGTTAGTATGTGACAGTACAAATGAGAGAGACTGTGGGCTTCTCTTCCCT 79
QY	27 LeuAlaIysGlnGlnValYsSerGlyThrIlySgIysSerGlnPrometSer----- 43
Db	80 GTGCTAGTATACCAATGGCTGTGTCGACACACACCAACCAATGAACCCAGCAATGCCATAC 139
QY	44 ProSerAspPheLeuAspIlySLeuMetGlyArgThrSerGlyTrysPalaArgIleArg 63
Db	140 GTCAAGACAGACATGACAGATTCCTCC-----AAGATATGACATTCGCTTCGCC 190
QY	64 ProAsnPhelYsGlyProProValAsnValThrCysAsnIlePheIleAsnSerPheSer 83
Db	191 CCGGACTTCGAGAGGCCCCCGCTGCACGTTGGGATGCGGATCGATGTCGCCAGCATAGAC 250
QY	84 SerValThrIysThrIleMetAspTrysArgValAsnValPheLeuArgGlnGlnTrpAsn 103
Db	251 ATGCTCTCGAAGTGAATGATGATATATACACGTACCACTGATATTCACAGACTTGGAAA 310
QY	104 AspProArgLeuSerTrysArgGlnTrysProAspSerLeuAspLeuAspProSerMet 123
Db	311 GACAAAGAGCCTTCTTATTCGTGAATCCCA---CTGAACCTCACCCCTAGACATAGAGTA 367
QY	124 LeuAspSerIleTrpIlysProAspLeuPhePheIleAsnGlnIlySgIyAlaAsnPheHis 143
Db	368 GCTGACCAACACTCGGGGTACACAGACACTTTCGATGACAAAGAAATCATTTGTGCAT 427
QY	144 GlnValThrIysAsnIlySLeuLeuAlaGlyIlePheIlyAsnGlyAsnValLeuTrpSer 163
Db	428 GGGCTCACTGAAAATGCAATGATTCGACTGCATCCTGATGACACAGTTCTTATGCA 487
QY	164 IleArgLeuThrIleLeuLeuSerCysLeuMetAspLeuIlyAsnPhePrometAspIle 183
Db	488 CTCGGATTCACCAACCAACAGCTGATGATGATGATTCGAAAGATATACACGTGATGAG 547
QY	184 GlnThrCysThrMetGlnLeuGlnIlySerSerIleLeuCysSerProLeuProSerLeu 203
Db	548 CAGAAAGCCACCCCTGGAGATCGAA----- 571
QY	204 SerLeuSerValGlyTrysIleMetIlySAspLeuValPheGlnTrpLeuGlnAspAlaPro 223
Db	572 -----AGTATAGGCTATACCACTGATGACATTTTACTGGAATGACGAGGAAGG 625
QY	224 AlaValGlnValAlaGlnGlyLeuThrLeuProGlnPheIleLeuArgAspGlnIlyAsp 243
Db	626 CAGCTACACTGCTGTATATAAATGCAACTTCCTCAATTTTCAATGTGTTGACTTACAAATG 685
QY	244 LeuGlyIlySgIysThrIlyStryrAsnThrGlyIlySphenIlyCysIleGlnValIlyPhe 265
Db	686 GTGCTAAGAGGTGGAG--TTCAACACAGGAGGCTATCCACAGCACTGTCACTAAGTTT 742
QY	264 HisLeuGlnArgIleMetGlyTrysIleLeuIleGlnMetIlyrIleProSerLeuIleIle 283
Db	743 CGTTAAAGAAACATGTGTTACTTCATTTTGCACAACTACATGCCCTTACACTGAT 802
QY	284 ValIleLeuSerTrysPheTrpIleAsnMetAspAlaIleProIleArgValGly 303
Db	803 ACAATTCCTGTCGAGGGGTCTTTTGATGACCACTATGATGATCTGCAAGCAAGTCCGA 862
QY	304 LeuGlyIleThrThrValLeuThrIleMetThrGlnIlySerSerGlySerArgAlaSerLeu 322

```

Db      863 CTAGAGATTCAGCAGCGTCTTACCAATGACACACCATCAGCACCACCTCAGGAGACCCG 92
Oy      324 ProLysValSerIerYrValLysAlaIleAspIleTrrMetAlaValCysLeuLeuPheVal 34
Db      923 CCAAAAGATCCCTATGATCCAAAGGATGGATATTTATCATGTGATGGCTGTGTTGTTGG 98
Oy      344 PheAlaIleLeuLeuGluTrrYrAlaIleAlaLeuAsnPheVal 356
Db      983 TTCCTGGCTCTGCTGGAGATGAGCTTTGTAATTTACATC 1021

RESULT 13
US-10-211-673-11
Sequence 11, Application US/10211673
Publication No. US20030013158A1
GENERAL INFORMATION:
APPLICANT: Le Boulanger, Beatrice
            Whiting, Paul John
TITLE OF INVENTION: HUMAN ALPHA 4 RECEPTOR SUBUNIT
                    OF THE GABA-A RECEPTOR
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.
City: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/211,673
FILING DATE: 02-Aug-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB95/02323
FILING DATE: 29-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME: Xu, Yang
REGISTRATION NUMBER: 45,243
REFERENCE/DOCKET NUMBER: T1292
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-1307
TELEFAX: 732-594-4720
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1555 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 47..1402
OTHER INFORMATION:
SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-10-211-673-11

Alignment Scores:
Pred. No.:
Score: 9,42e-77 length: 1555
Percent Similarity: 680,50 Matches: 161
Best Local Similarity: 50,11% Conservative: 63
Query Match: 36,02% Mismatches: 150
Db: 30,22% Indels: 75
Gaps: 7
US-10-075-846-4 (1-431) x US-10-211-673-11 (1-1555)
Oy      15 TrpThrLeuProGlyGlnValLeuLeuValGValAlaLeuAlaLysGluGluValLysSer 34
         ||||| ||| :||||| :|| :||

```

Db 48 TGGAGCGG -CCCCGGCGCTCTGGCCCCCGCTCGTCTCTGGCGGCACGACTCGG 106

Oy 35 GLyThrLySGlySerGlnProMetSer-----ProSerAspPhe 47

Db 107 GGCACCAAGCGATGAATGATCGTGGCGGACATACGCGGCTCCAACTGGATATCTCTGG 166

Oy 48 LeuAspLyLeuMetGlyArgThrSerGlyArgAspAlaArgIleArgProAsnPhelys 67

Db 167 CTCCCAACCTCGAGCGGGCTGATAGCCGGTTACGCCGCAACTTCGGCGCTGGCATCGGA 226

Oy 68 GlyProProValAsnValThrCysAsnIlePheIleAsnSerPheSerSerValThrLyS 87

Db 227 GGGCCCCCGCTGATATGTGGCCCTTGGCCCTGGAGGTGGCCAGCATGTGCACATCTCTCAAG 286

Oy 88 ThrThrMetAspTryArgValAsnValPheLeuArgGlnGlnItrPasnAspProArgLeu 107

Db 287 GCCAAACATGAGGTACACCATGACGGTGTCTCGCAACGACGTGGGGGAGCGAGCGTCTC 346

Oy 108 SerTryArgGlyLutryrProAspAspSerIleuAspLeuAspProSerMetLeuAspSerIle 127

Db 347 TCTCATC--AACCAACACCAACGAGACCTCTGGCGCTGGACACCGCTTGTGTGACAACTG 403

Oy 128 TrpLyPProAspPhePhePheAlaAsnIluGlyGlyAlaAsnPhenIleGluValPheThr 147

Db 404 TGGCTGCCCGACACCTTATGCTGTGACGCCCAAGTGCCTGGTGTCTACAGAGCTGCAGGTG 463

Oy 148 AspAsnLyLeuLeuArgIlePheLyAsnGlyAsnValIleuTrySerIleArgLeuThr 167

Db 464 GAGAAACAGCTCATCTCCGGCTGACGCCGACGCGGGGTGATCTGTACAGCATCCGAATCAC 523

Oy 168 LeuIleLeuSerCysLeuMetAspLeuLyAsnPhleProMetAspIleGlnThrCysThr 187

Db 524 TCCACTGGCGCTCGGACATGACCTGGCCAAATTCCTCCATGACGACGACGAGAGTGCATG 583

Oy 188 MetGlnLeuGlnSerSerSerIleuGlySerProLeuProSerLeuSerLeuSerVal 207

Db 584 CTGGACCTGGAG-----AGCTAC 601

Oy 208 GLyTryThrMetLyAspLeuValPheGluTrpLeuGluAspAlaProAlaValGlnVal 227

Db 602 GGTTACATCATCGGAGGACATCGTCTACACTACGTGGGAGACCAAGACACATCCACGG 661

Oy 228 AlaGluGlyLeuThrLeuProGlnIlePheIleLeuArgAspGluLyAspLeuGlyCys 247

Db 662 CTGGACAGCTGCAGCTGGCGCAGTTCACCATCACACAGCTACCGCTTCACACGAGAGCTG 721

Oy 248 ThrLyShiSTyrAsnThrGlyLyPheThrCysIleGluValLyPheHisLeuGluArg 267

Db 722 ATGAACCTTCAAGCTCCGCTGGCCAGTTCACCGGCTCACCGCTCCACTTCGGGGAGG 781

Oy 268 GlnMetGlyTryrTrpLeuIleGlnMetLyrlleProSerLeuLeuIleValIleLeuSer 287

Db 782 AACCGCGGCGGTATCATATCCAACTCCATCATGCTCCCTCGCTCTGTGGTCCATTGCTC 841

Oy 288 TrpValSerPheTrpIleAsnMetAspAlaIleProAlaArgValGlyLeuGlyIleThr 307

Db 842 TGGGCTCTCTCTGGATCAGCAGGCGGCGGCGCCGACAGGTTCTCTAGCATCAC 901

Oy 308 ThrValLeuThrMetThrThrGlnSerSerIleArgIleSerLeuProLySValSer 327

Db 902 ACGGTGCTGACATGACACCGCTCATGTGTCAAGTGCCTCTCTCCCTCCACCGGCACTCA 961

Oy 328 TyrValLySAlaIleAspIleTrpMetAlaValCysLeuLeuPheValPheAlaIleLeu 347

Db 962 GCCATCAAGGACATCGAGCTCTACTCTTCGATCTGCTATGTCTTGTGGTGGCCGCTG 1021

Oy 348 LeuGluTryrAlaAlaIleAsnPh----- 355

Db 1022 GTGGAGTACGCGCTTGTCTCATTTTCAACCGCACTACAGAGAAACGAAAGCCAAAGTCTC 1081

Oy 356 ---ValSerArgGlnHisLyGluPhe----- 363

Db 1082 AAGGTCTACAGGCGCGAGGAGATGGAGCTGAGAGAAACGCAATTGTCTCTTCTCCCTC 1141



```

OY 364 -----1leargleuargarglunargarglunargleu 375
DB 1142 TCTGTCGGCGGCGACAGAGAGCTGGCATCTCCCGCGACGGCGGTCCGGG- 1200
OY 376 GluGluaspIleIleGlnGluSerArgPheArgGlyTYRGLYLeuGlyHiscys 395
DB 1201 GAACCTGAT-----GGGCTCTACAGGTC 1224
OY 396 LeuGlnAlaArgaspGlyProMetGluGlySerGlyIleTYRserProGlnProPro 415
DB 1225 GGTGGCGGTGAGACAGAGGAGAGACGAGAGAGAGAGGGGCG- 1263
OY 416 AlaProLeuLeuArgGly 422
DB 1264 AGCCCGCTCAGAGGCGCAGG 1284

RESULT 14
US-10-239-420-13
; Sequence 13, Application US/10239420
; Publication No. US20030096984A1
; GENERAL INFORMATION:
; APPLICANT: Cully, Doris F.
; TITLE OF INVENTION: DNA MOLECULES ENCODING LIGAND GATED ION
; FILE REFERENCE: 20629P
; CURRENT APPLICATION NUMBER: US/10/239,420
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: PCT/US01/09956
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: 60/193,935
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 1150
; TYPE: DNA
; ORGANISM: Dermacentor variabilis
US-10-239-420-13

Alignment Scores:
Pred. No.: 1.26e-75 Length: 1150
Score: 670.00 Matches: 130
Percent Similarity: 64.24% Conservative: 73
Best Local Similarity: 41.14% Mismatches: 91
Query Match: 29.75% Indels: 22
DB: 9 Gaps: 7

US-10-075-846-4 (1-431) x US-10-239-420-13 (1-1150)
OY 46 AspheLeuaspLysLeuMetGlyArgThrSerGlyTYRaspAlaArgIleArgPro--- 64
DB 83 GACATTCGAGCAGAGCTCTC-----AAAACTACGATCGAAGGCCCTGCGCAGC 133
OY 65 AsnPhelysGlyProProValAsnValThrCysAsnIlePheIleAsnSerPheSer 84
DB 134 AGTCACCTCGGAATGCAACTATTGTCTGCATGCGAAATTTACATACGAAGTTTGATCA 193
OY 85 ValThrIysThrThrMetAspTYRArgValAsnValPheLeuArgGlnGlnITTPAsnSP 104
DB 194 ATAAATCTTGCAACATGACTACGACGAACTCACTTACCTTCGCGCACTGCTGCTGCAC 253
OY 105 ProArgLeuSerTYRArgGlyTYRProAspSerLeuAspLeu---AspProSerMet 123
DB 254 GAGCGGTACCC---AAATCACGCTATCTCTGCTCGCTCGACCTTAATGACCCAAAGCTG 310
OY 124 LeuaspSerIleTYRPLysProAspLeuPhePheAlaAsnGlyLysGlyAlaAsnPhenH 143
DB 311 GTACAAATGATATGAGAGCCAGAGATTCTTCTTGCAAGCGGAAACAGCGGAGTTCCAA 370
OY 144 GluValThrThrAspAsnLysLeuLeuArgIlePheIysAsnGlyAsnValLeuTYRser 163

```

```

DB 371 TATGACTGTACCTAACGCTCTGTTAGATCAACCCGAGTGAATATCTTTACATG 430
OY 164 IleargleuThrIleLeuSerCysLeuMetAspLeuLysAsnPhProMetAspIle 183
DB 431 TTGCGTTAAACTAGAGTCTCTGCTGATGAGCTGTACCGGTACCCATGATTC 490
OY 184 GlnThrCysThrMetGlnLeuGluSerSerIleLeuCysSerProLeuProSerLeu 203
DB 491 CAAGTCTCAGCATCGAATAATGCCCTTTTCC----- 523
OY 204 SerLeuSerValGlyTYRThrMetLysAspLeuValPheGluITTPLeuGlnAlaPro 223
DB 524 -----AAACACCGCAAGAGCGCTGCTGAATATGTCGACAGTCAGCT 568
OY 224 AlaValGlnValAlaGluGlyLeuThrLeuProGlnPheIleLeuArgaspGlyLysasp 243
DB 569 GTCGTT---CTCTTCGATTAACCTCAAGTGGCCCACTTTGAATA---GACAAAGTGAC 622
OY 244 LeuGlyCysCysThrLysHisTYRAsnThrGlyLysPheThrCysIleGluValLysPhe 263
DB 623 ACGTCCTTATGCAAGAAGAAAGTTTCACATAGGGGAATACAGTTGCTGAAGCCGACTTC 682
OY 264 HisLeuGluArgGlnMetGlyTYRTYRLeuIleGlnMetTYRLeuProSerLeuLeuIle 283
DB 683 TATTCGACGCTTCCTCGTTATCACATGGTCGACAGCTATCTTCCACACCGCTTATC 742
OY 284 ValIleLeuSerTYRProValSerPheThrPheIleAsnMetAspAlaAlaProAlaArgValGly 303
DB 743 GTGTCATCTATGAGGTCTCATTTCTGCTCAGCAGAGCGCAATACCGCGGTGTACC 802
OY 304 LeuGlyIleThrThrValLeuThrMetThrThrGlnSerSerGlySerArgLysSerLeu 323
DB 803 CTGGCGCTAACACAGCGCTCTCACCATCTCATTCAAGAGGTCGGGTATCCAGGAACCTG 862
OY 324 ProLysValSerTYRValLysAlaIleAspIleThrPheMetAlaValCysLeuLeuPheVal 343
DB 863 CTCTCCGCTCTGCTCATCAAGAGCGCATCTGGAATGAGATCTGACTTCTTTGTC 922
OY 344 PheAlaAlaLeuLeuGluTYRAlaAlaIleAsnPheValSerArgLys 359
DB 923 TTTGGCGCCCTTCTAAGATTCACTTCATCAATATCTCTGAGGCGG 970

RESULT 15
US-10-239-420-4
; Sequence 4, Application US/10239420
; Publication No. US20030096984A1
; GENERAL INFORMATION:
; APPLICANT: Cully, Doris F.
; TITLE OF INVENTION: DNA MOLECULES ENCODING LIGAND GATED ION
; FILE REFERENCE: 20629P
; CURRENT APPLICATION NUMBER: US/10/239,420
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: PCT/US01/09956
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: 60/193,935
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 2194
; TYPE: DNA
; ORGANISM: Dermacentor variabilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (47) ... (1315)
US-10-239-420-4

Alignment Scores:
Pred. No.: 3.16e-73 Length: 2194
Score: 655.00 Matches: 158
Percent Similarity: 52.53% Conservative: 91

```

Best Local Similarity: 33.33% Mismatches: 133  
Query Match: 29.09% Indels: 94  
DB: 9 Gaps: 14  
US-10-075-846-4 (1-431) x US-10-239-420-4 (1-2194)

QY 18 ProGlyInValLeuLeuArgValAlaLeuAlaLysGluValLysSerGlyThrLys 37  
Db 123 CCTGGTTCGCTCAGGCTGTGCACACATCGACAAAGAT-----ACCCAGACACAG 173  
QY 38 GlySerInProMetSerProSerAspPheLeuAspLysLeuMetGlyArgThrSerGly 57  
Db 174 GACCTTCGATGATGATGACCTT-CAATATTTGGACACACATCTTATGG-----AGC 223  
QY 58 TyrAspAlaArgIleArgPro---AsnPhelysGlyProProValAsnValThrCysAsn 76  
Db 224 TACACCCGACCGCATCACCCTCGGTATCATTTAAACGTTCTACAGTTGTTAAAGTGGCAG 283  
QY 77 IlePheIleAsnSerPheSerSerValThrLysThrMetAspTyrArgValAsnVal 96  
Db 284 ATATATCTCAGAGATTGTTGGAGCTGTGAACCTCACAACATGACTAGACGTAGACCTG 343  
QY 97 PheLeuArgGlnGlnTTPAsnAspProArgLeuSerTyrArgGluTyrProAspAspSer 116  
Db 344 TACCTGGCTCAGACGTGACGACTTGGCGATGAAGAACCCCAACCTG--ACCCGCTCC 400  
QY 117 LeuAspLeu---AspProSerMetLeuAspSerIleThrLysProAspLeuPhePheAla 135  
Db 401 CTAGACTTTAAAGCAGCCCAACCTCTCAGAAAGTGTGGAACCTGACGCTTACTTTCCC 460  
QY 136 AsnGluLysGluAlaAsnPheHisGluValThrThrAspAsnLysLeuLeuArgIlePhe 155  
Db 461 AATGCCAAGCAGCGGAGGATTCAGTCTGCTACGTTCCCAACGTTCTTGGATGATATAC 520  
QY 156 LysAsnGluAsnValLeuTyrSerIleArgLeuThrLeuIleLeuSerCysLeuMetAsp 175  
Db 521 CCTACCGCGCATATACCTTACATGTTAAGGCTTAAGCATCTCTCGCATGATGAAAC 580  
QY 176 LeuLysAsnPheProMetAspIleGlnThrCysThrMetGlnLeuGluSerSerIle 195  
Db 581 ATGGAGCGGTACCCCTGGACCGACAGGTCTGCAGCATCGACTTGGCTTCTTTTCC--- 637  
QY 196 LeuCysSerProLeuProSerLeuSerLeuSerValGlyTyrThrMetLysAspLeuVal 215  
Db 638 -----AAGACGACAAAGAGGTTGAG 658  
QY 216 PheGluTrpLeuGluAspAlaProAlaValGlnValAlaGluGluLeuThrLeuProGln 235  
Db 659 CTCCAATGG--GGAAAGCTGAGAGCTGTGCACATGTACAGTGGCTGGAAGATGACACA 715  
QY 236 PheIleLeuArgAspGluLysAspLeuGluCysCysThrLysHisTyrAsnThrGlyLys 255  
Db 716 TTTCAGCTTCA---CAATACAGCTGACGAGATGCAAGGCGGCTTTCAATAGAGCGAG 772  
QY 256 PheThrCysIleGluValLysPheHisLeuGluArgGlnMetGlyTyrThrLeuIleGln 275  
Db 773 TACAGCTGCTGCGCGGAGCTCAACTTAAGCGTTCCATTGGCGACACACACATGATGAG 832  
QY 276 MetTyrIleProSerLeuLeuIleValIleLeuSerTrpValSerPheTrpIleAsnMet 295  
Db 833 TCTTACCTGCGCTCACATCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 892  
QY 296 AspAlaAlaProAlaArgValGlyLeuGlyIleThrThrValLeuThrMetThrThrGln 315  
Db 893 GACCGCATACCGGCGGATACCGCTGGGTGTACACACGCTCTCCTATTTCTGCGGAG 952  
QY 316 SerSerGlySerArgAlaSerLeuProLysValSerTyrValLysAlaIleAspIleTrp 335  
Db 953 AGCTCCGACACACAGGCAACCTAGCGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1012  
QY 336 MetAlaValLysLeuLeuPheValPheAlaAlaLeuLeuGluTyrAlaAlaIleAsnPhe 355  
Db 1013 ATGGGACAGCTGCACCATGT 1072

QY 356 ValSerArgGlnHisLysGluPheIleArgLeuArgArgGlnArg-----Arg 372  
Db 1073 CTCGCTGCAAGAA--GCAGATCGTGGCCGCTCATGTGCGGACGTGACGCTTCCCAAGA 1131  
QY 373 GlnArg-----LeuGluGluAsp 378  
Db 1132 TCTGTTCTTGTCTGTGGGAAACAGACAAATAATGCACCCCGTCAACGTTCCCGGAC 1191  
QY 379 Ile-----IleGlnGlu 382  
Db 1192 GTCCACCCAGT 1251  
QY 383 SerArgPheTyrPheArg----- 388  
Db 1252 GGTGCGTTTCCCATCGGCTTGTTCCTTCAACGCACTGACGGCCCTATTACTTGTCT 1311  
QY 389 -----GlyTyrGlyLeu----- 392  
Db 1312 CTAGTTGGCCATGTGTCTAGTGTCTACAGACTGTCTCCCAACGTGACGCCATGACCGG 1371  
QY 393 -----GlyHisCysLeuGlnAlaArgAspGlyIlePheMetGluGlySerGlyIleTyr 410  
Db 1372 GAAACGGGTGGCTGTACATCCCAAGGAAACGTCGCGCGGTGTGAATGAAGAAAGACTG-- 1428  
QY 411 SerProGlnProProAlaProLeuLeuArgGluGlyGluThr 424  
Db 1429 ---CATTCACCGACGACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1467

Search completed: July 1, 2003, 01:01:34  
Job time : 248.481 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

# OM protein - nucleic search, using frame\_plus.p2n model

Run on: June 30, 2003, 23:00:29 ; Search time 2265.88 seconds  
(without alignments)  
3080.588 Million cell updates/sec

Title: US-10-075-846-4  
Perfect score: 2252  
Sequence: 1 MTTVPATLSFLTLMTLPQG.....PQPPAPLREGETTRKLYVD 431

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n model -DEV=xlp  
-O=/cg2.1/USPTO\_spool/US10075846/runat\_25062003\_163648\_5157/app\_query.fasta\_1.782  
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human4.0.cdi -LIST=45  
-DOALIGN=200 -THR.SCORE=Pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pco -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10075846.ecgn.1.1.3724.etunal.25062003.163648.5157 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOC  
-REV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
EST :  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estlmu:\*  
5: em\_estlov:\*  
6: em\_estlpl:\*  
7: em\_estlro:\*  
8: em\_hlc:\*  
9: gb\_estcl:\*  
10: gb\_estl2:\*  
11: gb\_hlc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estum:\*  
16: em\_estum:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_tod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1033.5	45.9	662	9	AUI69868
2	877.5	39.0	779	14	BQ443713
3	854.5	37.9	678	10	AV729257
4	853.5	37.9	2991	11	BC027094
5	832.5	37.0	2593	11	BC022502
6	789.5	35.1	692	12	BE981841
7	772.5	34.3	720	12	BE981459
8	746.5	33.1	817	12	BG404477
9	672.5	29.9	975	14	BQ938794
10	631.5	28.0	2475	11	AK013727
11	631	28.0	893	9	AL538200
12	619.5	27.5	802	14	BQ042344
13	611	27.1	817	9	AUI33223
14	605	26.9	1942	11	BC029850
15	570.5	25.3	948	9	AL539373
16	555	24.6	3257	11	BC026337
17	548.5	24.4	536	12	BF906462
18	548	24.3	552	13	BM440664
19	545	24.2	957	9	AL572853
20	519	23.0	586	10	AM122941
21	510	22.6	594	12	BF937654
22	504	22.4	2708	11	AK018768
23	501	22.2	690	14	BQ044592
24	499	22.2	723	13	B1736340
25	499	22.2	908	13	B1825190
26	498.5	22.1	663	10	BB653397
27	497	22.1	831	13	B1663922
28	488.5	21.7	1134	14	BQ667322
29	488	21.4	1068	13	BM547796
30	482.5	21.4	623	10	AM077068
31	479	21.3	1078	13	B1757807
32	470	20.9	785	14	BM964113
33	469.5	20.8	843	13	B1913857
34	465	20.6	1047	14	BQ067799
35	461.5	20.5	501	10	AM280976
36	460.5	20.4	912	13	B1829118
37	460.5	20.4	975	13	B1552703
38	460	20.4	867	12	BG189196
39	460	20.4	918	13	B1824761
40	459.5	20.4	582	13	B1681394
41	457	20.3	567	13	BM129587
42	457	20.3	570	13	BM129116
43	456	20.2	814	12	BG706397
44	446.5	19.8	861	13	B1553004
45	445.5	19.8	700	12	BG404089

## ALIGNMENTS

RESULT 1  
AUI69868  
LOCUS AUI69868 662 bp mRNA linear EST 29-JAN-2001  
DEFINITION AUI69868 Ol-br-ad cDNA Oryzias latipes cDNA clone br5332, mRNA  
ACCESSION AUI69868  
VERSION AUI69868.1 GI:12591937  
KEYWORDS EST.  
SOURCE Japanese medaka.  
ORGANISM Oryzias latipes  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
Acanthomorpha; Acanthopterygii; Percomorpha; Albetinomorpha;  
Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.  
REFERENCE 1 (bases 1 to 662)

AUTHORS Mita, K., Ishikawa, Y. and Yamauchi, M.  
 TITLE Establishment of cDNA database of medaka, *Oryzias latipes*  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Mita, K

Genome Research Group  
 National Institute of Radiological Sciences  
 Anagawa 4-9-1, Inage, Chiba 263-8555, Japan  
 Email: kmita@nirs.go.jp  
 method: uni-directional sequence direction: sequenced from T3 primer  
 (5' -> 3').

# FEATURES

source Location/Qualifiers  
 1..662  
 /organism="Oryzias latipes"  
 /strain="HN1"  
 /db\_xref="taxon:8090"  
 /clone="pr5332"  
 /clone\_1lb="OI-br-ad cDNA"  
 /sex="female/male mixed"  
 /tissue\_type="brain"  
 /dev\_stage="adult"  
 BASE COUNT 155 a 205 c 164 g 138 t

## ALIGNMENT SCORES:

Pred. No.: 1,66e-121 Length: 662  
 Score: 1033.50 Matches: 193  
 Percent Similarity: 88.09% Conservative: 14  
 Best Local Similarity: 82.13% Mismatches: 13  
 Query Match: 45.89% Indels: 15  
 DB: 9 Gaps: 2

US-10-075-846-4 (1-431) x AU169868 (1-662)

OY 66 PhelysGlyProProValAsnValThrCysAsnIlePheIleXnsrPheSerVal 85  
 |||||  
 DB 2 TTTAAAGGTCCACCTTAACGCGACCTGCACATTTCATCAGCGCTTTGGATCCATC 61  
 |||||  
 OY 86 ThrLysThrThMetAspTyrArgValAsnValPheLeuArgGlnGlnTyrPasnPro 105  
 :::::  
 DB 62 GCGAAGACCAACCATGAGTACAGAGTGAACATCTCTGAGCGAGCATGGAACGCC 121  
 |||||  
 OY 106 ArgLeuSerTyrArgLysTyrProAspAspSerLeuAspLeuAspProSerMetLeuAsp 125  
 |||||  
 DB 122 CGCGTGGCTACAGAGTATCCGCGACTCGCTGATCTGGACCGCTCATGTTGGAC 181  
 |||||  
 OY 126 SerIleTyrPyrAspLeuPhePheAlaAsnGluLysGlyValAsnPheHisGluVal 145  
 |||||  
 DB 182 TCCATCTGAGAGCGCATCTGTTCTTCTTAATGAGAAGGGGCCAATCTCCACGAGTTC 241  
 |||||  
 OY 146 ThrThrAspAsnLysLeuLeuArgIlePheLysAsnGlyAsnValLeuTyrSerIleArg 165  
 |||||  
 DB 242 ACCACGAGCAACAGCTGCTGCGCATCTCCAAAATGAGCAAGCTGCTATACACATACGA 301  
 |||||  
 OY 166 LeuThrLeuIleLeuSerCysLeuMetAspLeuLysAsnPheProMetAspIleGlnThr 185  
 :::::  
 DB 302 ATCACTCTATCTCGCTGCGCTCCCATGATCTGMAAAGCTTCCCATGATGTGCAGACC 361  
 |||||  
 OY 186 CysThrMetGlnLeuLysSerSerIleLeuCysSerProLeuProSerLeuSerLeu 205  
 |||||  
 DB 362 TGCATCATCGACGTGAG----- 379  
 |||||  
 OY 206 SerValGlyTyrThrMetLysAspLeuValPheGluTyrLeuGluAspAlaProAlaVal 225  
 |||||  
 DB 380 AGCTTCGCTACACCATGAGACGCTCATCTTCGACTGG--GACAGAGAAGGGCGCGTGG 436  
 |||||  
 OY 226 GlnValAlaGluGlyLeuThrLeuProGlnPheIleLeuArgAspLysAspLeuGly 245  
 |||||  
 DB 437 CAGGTGGCGAGCGCTGAGCGCTCGCTCATCTCCTAAGAGAGAGAGAGAGAGAGAGAG 496  
 |||||  
 OY 246 CysCysThrLysHisTyrAsnThrGlyLysPheThrCysIleGluValLysPheHisLeu 265  
 |||||  
 DB 497 TACTGACCAAGACACTACACACAGGTAAATTCACCTGATGAGGCTTCACACCTGG 556

OY 266 GluArgGlnMetGlyTyrTyrLeuIleGlnMetTyrIleProSerLeuLeuIleValIle 285  
 |||||  
 DB 557 GAGCGACACATGGCGTACTACTATCCAGATGATACATCCCTCGCTGATCGTCAATC 616

OY 286 LeuSerTyrValSerPheThrPheLeuAsnMetAspAlaIleProAla 300  
 |||||  
 DB 617 CTGCTCGGTGCTCTCTTGATGATCAACATGAGAGCGCGCGCGCC 661

RESULT 2  
 BQ443713 779 bp mRNA linear EST 29-MAY-2002

LOCUS BQ443713  
 DEFINITION UI-M-EMO-bxh-j-02-0-UI.r1 NIH\_BMAP\_EMO Mus musculus cDNA clone  
 IMAGE:5708257 5', mRNA sequence.

ACCESSION BQ443713  
 VERSION BQ443713.1 GI:21246825

KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus

REFERENCE NIH-MGC http://mgi.nci.nih.gov/  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 TITLE Unpublished (1999)  
 JOURNAL Contact: Robert Strausberg, Ph.D.  
 COMMENT Email: cga@bbs.femail.nih.gov  
 Tissue Procurement: Dr. James Lin, University of Iowa

cDNA library preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA library arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov

This clone was contributed by the Brain Molecular Anatomy Project  
 (BMAP)

Seq primer: pyx-5.

## FEATURES

source

1..779

Location/Qualifiers

/organism="Mus musculus"

/strain="C57BL/6"

/db\_xref="taxon:10090"

/clone="IMAGE:5708257"

/clone\_1lb="NIH\_BMAP\_EMO"

/tissue\_type="whole brain"

/dev\_stage="embryo 15.5 dpc"

/lab\_host="DH10B (T1 phage resistant)"

/note="Organ: brain; Vector: pyx-asc; Site\_1: EcoR I; Site\_2: Not I; The library was constructed according to Bonaldi, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pyx-asc vector. The library tag sequence located between the Not I site and the polyA tail is GTCCGTCGAA. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP). 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

BASE COUNT 222 a 167 c 165 g 225 t

## ALIGNMENT SCORES:

Pred. No.: 2.52e-101 Length: 779  
 Score: 877.50 Matches: 174  
 Percent Similarity: 78.99% Conservative: 14  
 Best Local Similarity: 73.11% Mismatches: 29  
 Query Match: 38.97% Indels: 21  
 DB: 14 Gaps: 4

US-10-075-846-4 (1-431) x BQ443713 (1-779)







```

DEFINITION Homo sapiens, glycine receptor, beta, clone IMAGE:4792516, mRNA.
ACCESSION BC022502
VERSION BC022502.1 GI:18490294
KEYWORDS htc.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 2593)
JOURNAL Strausberg, R.
Direct Submission
Submitted (01-FEB-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.ncl.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Miklos Pavlovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshituki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www.shgc.stanford.edu
Contact: (Dickson, Mark) mdc@paxill.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 32 Row: D Column: 19
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 4504022
This clone has the following problem: frame shifted.

FEATURES
Source
location/Qualifiers
1..2593
/organism="Homo sapiens"
/db_xref="LocusID:2743"
/db_xref="taxon:9606"
/clone="IMAGE:4792516"
/tissue_type="Brain, hypothalamus"
/clone_lib="NH_MGC_96"
/lab_host="DH10B"
/note="vector: pBluescript"
BASE COUNT 824 a 466 c 488 g 815 t
ORIGIN
Alignment Scores:
Pred. No.: 1.54e-94 Length: 2593
Score: 832.50 Matches: 186
Percent Similarity: 63.78% Conservative: 64
Best Local Similarity: 47.45% Mismatches: 89
Query Match: 36.97% Indels: 54
DB: 11 Gaps: 11

US-10-075-846-4 (1-431) x BC022502 (1-2593)
Oy 8 ThrLeuSerPheLeuLeuTrpPheProGlyGlnValLeuLeuArgValAlaLeu 27
||| ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 116 ACAACGCGCTTTTAAATTATTTAAATTTCTTG-----TAAGTGAAGAAGACCTAT 163
Oy 28 AlaLysGlnGlnValLysSerGlyThrLysGlySerGlnPromet----- 42
:::|||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 164 TCTAAGGAAAAGTCTTCAAGAAGGGAAGGAAAAAAGAAGACGATATCTGCCCATCT 223
43 -----SerProSerPheLeuasp 49
224 CAGCAGTCAGCAGAGACCTTGCCCGAGTACCTGCCAACCTCAGCAGCAATATCTTGAC 283
Oy 50 LysLeuMetGlyArgThrSerGlyTyrAspAlaArgGlnLeuGpProAsnPheLysGlyPro 69
:::|||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 284 AGGTTATTG-----GTCAGATTATATCCCAAGATTAAGCCAAACTTCAAGGCAATT 334

```

Oy	70	ProValaSnValThrCysAsnIlePheIleAsnSerPheSerValThrLysThr	89
Db	335	CCGTGTAATGAGTAGTACACATTATTTATTAACAGTTTGGATTCATTCAAGAACAACA	394
Oy	90	MetAspTYrArgValIaSnValPheLeuAArgGlnIleTrrPaSnAspProAArgLeu	107
Db	395	ATGAGACATATAGAGTTAAACATCTTCTCGAGACAAAATGGAAATGACCCAGGTTCAACTC	454
Oy	108	-----SerTYrArgGluTYrProAspAspSerIleuAspIleuAspProSerMetLeuAsp	125
Db	455	CCCAAGTATTTTAGG-----GGTTCAAGATGCACGTGACAGGCGATCCACAAATCTCAAG	508
Oy	126	SerIleTrrLysProAspLeuPhePheAlaSnGlnIuLysGluValAsnPheHisGluVal	145
Db	509	TGTTTATGGAACACTGATTTATTTTTCGCAATATAAAAAGTGGCAATTTTTCATGATGTG	568
Oy	146	ThrThraSpaSnLysLeuLeuAArgIlePheLysAsnGlyAsnValLeuTYrSerIleArg	165
Db	569	ACCCAGCAAAACATCCCTCCCTTATTTTTCGTATGATGAGATGCTTCCTTGACATGATG	628
Oy	166	LeuThrIleuIleLeuSerCysLeuMetCAspLeuLysAsnPheProMetAspIleGlnTrr	185
Db	629	TTATCTATTACTCTTTTATGCGCTTTGGACTTGCATTTGTTTCCAAATGGATACCAACGT	688
Oy	186	CysThrMetGlnLeuGlnSerSerSerIleuLysSerProLeuProSerLeuSerLeu	205
Db	689	TGCAAGATGCACACTGGAC-----	706
Oy	206	SerValGlyTYrThrMetLysAspLeuValPheGluTrrPLeuGlnAspAlaProAlaVal	225
Db	707	AGCTTTGGTTCACACACATGATGATTTACGATTTATCTGCGACGACGAGATCCT---GTG	763
Oy	226	GlnValAlaGluGlyLeuThrLeuProGlnPhe---IleLeuAArgAspGluLysAspLeu	244
Db	764	CAATTTA---GAAAAAATTTGCCCTTCCTCAATTTATATCAAAAAGGAAGATATTGAATAT	820
Oy	245	GlyCysCysThrLysHisLysTYr---AsnThrGlyLysPheThrCysIleGluValLysPhe	263
Db	821	GGTAACACTGTACAAATACTATAAAGCAGCGGCTACATACATCGCTGGAGAAAGTATCTTC	880
Oy	264	HisLeuGlnAArgGlnMetGlyTYrTYrTYrLeuIleGlnMetTYrIleProSerLeuLeuIle	283
Db	881	ACCTTGAGGAGGACGAGGCTGCTTACATGATGGGGGCTCTGCGCCCAACCTCTCATTT	940
Oy	284	ValIleLeuSerTrrPValSerPheTrrPLeuMetLeuAspAlaIleProAlaArgValGly	303
Db	941	GTGTGTCTCTCTGCTTCTCTTCTGTGATCAACCGGACGCGAATGCTGCCACAGTGGC-	999
Oy	304	LeuGlyIleThrTrrValLeuThrMetThrGlnSerSerGlySerArgAlaSerLeu	323
Db	1000	CTGGGTATCTTCCACAGCTCAGCTCAGCTTGGCCCTGATGACACAACCTTTGCCGCTGACCTT	1059
Oy	324	ProLysValSerTYrValLysAlaIleAspIleTrrPheAlaValCysLeuLeuPheVal	343
Db	1060	CCCAAAAGTTTCTTATGTGTAAAGGCTCTGTAGCTTTATGCTTCTCTCTCTTGGC	1119
Oy	344	PheAlaIleLeuLeuGlnTYrValAlaIle-----AsnPheValSerArgGln	359
Db	1120	TTTCTCTTCCCTGGGGAGTATGACGTTTCCAGCTGATGCTGAACAACCCCAAAAGGCTT	1179
Oy	360	HisLysGluPheIleArgLeuArgArgArgGlnArg	371
Db	1180	GAAGCTGAAAAAAGCAGAAATTGCTAAAGGCTGACGA	1215
RESULT 6			
LOCUS	BE981841	692 bp	mRNA
DEFINITION	UI-M-CG0P-bdd-h-07-0-UI.s1 NIH BMAP_Ret4_s2 Mus musculus cDNA clone		
ACCESSION	BE981841		
VERSION	BE981841.1	GI:10651376	
KEYWORDS	EST.		



```

DEFINITION Homo sapiens, glycine receptor, beta, clone IMAGE:4792516, mRNA.
ACCESSION BC022502
VERSION BC022502.1 GI:18490294
KEYWORDS HTC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 2593)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (01-FEB-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: gcgaps-remail.nih.gov
Tissue Procurement: Miklos Pavkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www.shgc.stanford.edu
Contact: (Dickson, Mark) mcdpaxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
Series: IRAC Plate: 32 Row: b Column: 19
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 4504022
This clone has the following problem: frame shifted.
FEATURES
Source
1..2593
location/qualifiers
/organism="Homo sapiens"
/db_xref="LOCUSID:2743"
/db_xref="Caxon:9606"
/clone="IMAGE:4792516"
/tissue_type="Brain, hypothalamus"
/clone_lib="NH_MGC_96"
/lab_host="DH10B"
/note="vector: pBluescript"
BASE COUNT 824 a 466 c 488 g 815 t
ORIGIN
Alignment Scores:
Pred. No.: 1.54e-94 Length: 2593
Score: 832.50 Matches: 186
Best Local Similarity: 63.78% Conservative: 64
Percent Similarity: 47.45% Mismatches: 89
Query Match: 36.97% Indels: 54
DB: 11 Gaps: 11
US-10-075-846-4 (1-431) x BC022502 (1-2593)
Oy 8 ThrLeuSerPheLeuLeuLeuTriphLeuProGlyGlnValLeuLeuArgValAlaLeu 27
Db 116 ACAATTGCCCTTTTAAATTTTAAATTTCCCTTG-----TAAGTGAAGAAACCTTAT 163
Oy 28 AAlayGcIuGlnValUallySerGlyThrlyrGlySerGlnPromet----- 42
Db 164 TCTAAGGAAAAGCTTCAAGAAAGGGAAGGGAAGAAAGAACATATATGCCCATCT 223
Oy 43 -----SerProSerAsPheLeuAsp 49
Db 224 CAGCAGTCAGCAGACGACCTTGCCCGATACCTGCCAACATCCACATACATCTTGAAC 283
Oy 50 LysLeuMetGlyArgThrSerGlyIyrAspAlaArgIleArgProAsnPhelGlySlyPro 69
Db 284 AGGTAATATG-----GTCAAGTTATATCCCAAGATPAAGCAACAACTTCAAGGCATT 334

```

OY	70	ProvalasnaValmThrcysAsnIlePheIleuInsSerPheSerValmThrllysThrThr	89
Db	335	CCCTTTGATGAGTGAAGTCAACATTTTATTATACAGTTTGGATCTCATTTCAAGAACACACA	394
OY	90	MetAspTyrValArgValAsnValPheLeuArgGlnInlTrpAsnAspProArgLeu	107
Db	395	ATGGACATATAGATTACATCTTCTCTGACGACAAAATGGATGACCCAGGTTGAACCTC	454
OY	108	-----SerTyrArgGluTyrTrpAspSerSerLeuAspLeuAspProSerMetLeuAsp	125
Db	455	CCCAAGTATTTTAGG-----GGTTTCAGATGCATGACAGTGCATCCAAACATGTACAAAG	508
OY	126	SerIleTrpIlyProAspLeuPhePheAlaAsnGlnIlyuSglValAlaAsnPheHisGluVal	145
Db	509	TGTTTATGGAACCTCGATTATTTTATTTTGGCAATGAAAAAAGCTGCCAATTTTCATGATGTG	568
OY	146	ThrTrpAspAsnLysLeuLeuArgIlePheLysAsnGlnAsnValLeuTyrSerIleArg	165
Db	569	ACCCAGGAACATCCTCTCTCTTATTATTTTGGTATGACAGATGTCCTTGTCCAGATGAG	628
OY	166	LeuThrIleuIleLeuSerCysLeuMetAspLeuLysAsnPheProMetAspIleGlnThr	185
Db	629	TTATCTATTACTCTTTTCATGCTCCCTTGGACTTGCATGATTTTCCAAATGGATACACACGT	688
OY	186	CysThrMetGlnLeuGlnSerSerSerIleLeuCysSerProLeuProSerLeuSerLeu	205
Db	689	TGCAGATGACACATCGAC-----	706
OY	206	SerValGlyTyrTrpMetLysAspLeuValPheGluTrpLeuGlnAspAlaProAlaVal	225
Db	707	AGCTTTGGTTACACACATGATGATTTACGATTTATTCGCGACGACGAGATCCG---GTG	763
OY	226	GlnValAlaGluGlyLeuThrLeuProGlnPhe---IleLeuArgAspGluLysAspLeu	244
Db	764	CAATTA---GAAAAAATTCCTCTGCTCCCAATTGATACAAAAGGAAGATTTGAATAT	820
OY	245	GlyCysCysThrLysHisIstYr---AsnThrGlyLysPheThrCysIleGlnValLysPhe	263
Db	821	GGTAACTATACAAAATACTATATAAGGACCGGGCTACTACATGCGTGGAAAGTCATCTTC	880
OY	264	HisLeuGlnArgGlnMetGlyTyrTyrLeuIleGlnMetYrIleProSerLeuLeuIle	283
Db	881	ACCTGAGGAGGACGAGTGGCTTTTACATGATGGGGGTCTACGCCCAACCTGCTCATTT	940
OY	284	ValIleLeuSerTrpValSerPheTrpIleAsnMetAspAlaAlaProAlaArgValGly	303
Db	941	GTTGTTCTCTCTGCTCTCTCTTCCTTCGATCAACCGGACGGAGTGCACAGAGTGC-	999
OY	304	LeuGlyIleThrThrValLeuThrMetThrThrGlnSerSerGlySerArgAlaSerLeu	323
Db	1000	CTGGGATATCTCTCAGTCTCAGCTGAGCTGGCCCTCTGATGCACACACCCTTGGCCGTAGCTT	1055
OY	324	ProLysValSerTyrValLysAlaIleAspIleTrpMetAlaValCysLeuLeuPheVal	343
Db	1060	CCCAAGGTTTCTCTATGTAAGGCTCTTATTGGCTTATTGCTGCTCTCTCTTTGGG	1119
OY	344	PheAlaAlaLeuLeuGlnTyrAlaAlaIle-----AsnPheValSerArgGln	359
Db	1120	TTTGGTTCCTCCGTGGAGATATGACATTTGCCAGTGATGCTGAACACCCCAAAAGGGTT	1179
OY	360	HisLysGluPheIleArgLeuArgArgArgGln	371
Db	1180	GAAGCTGAAAAAGCCAGAAATTGCTTAAGCTGACGA	1215
RESULT 6			
LOCUS	BE981841	692 bp	mRNA
DEFINITION	UI-M-CG0P-bdd-h-07-0-UI-s1 NIH BMP Ret4.S2 Mus musculus cDNA clone		
ACCESSION	BE981841		
VERSION	BE981841.1	GI:10651376	
KEYWORDS	EST.		





```

OY 300 AlaArgValGlyLeuGlyIleThrValLeuThrThrThrGlnSerSerGlySer 319
Db 313 GCCAGTTGCC-CYGGCATTACAAACAGCTCCTGACATGACTACAGAGTTCCAGTTCC 371
OY 320 ArgAlaSerLeuPheValSerTyrValLysAlaIleAspIleTyrPmeAlaValCys 339
Db 372 AGGCA-TCCTGCGCAAGAGTCTCTTAAAGCAATGACATCTGATGGTGGTATGC 430
OY 340 LeuLeuPheValPheAlaAlaLeuLeuGlyTyrAlaAlaIleAsnPheValSerArgL 359
Db 431 CTCTCTTTGTGTGGTGGCTTACTGGAATATCAGACAGTGAAGCTT-CTCTCCAGGCAA 489
OY 360 HisLysGluPheIleArgLeuArgArgGlnArgGlnArgGlnArgGlnArgGln 379
Db 490 CATAGAGAGTTCCTCTCTCCGAGACGACAGAGAGGCAATTAAGAAGATGTT 549
OY 380 IleGlnIleSerArgPheTyrPheArgGlyTyrGlyLeuGlyHisCysLeuGlnAlaArg 399
Db 550 ACTGCTAAAGTCTTTTAACCTTACGATGGCTATGGATGGGTCTGCTCCAAATGAA 609
OY 400 AspGlyLeuPheMetGlySerGlyIleTyrSer--ProGlnProPro 415
Db 610 GATGGACAGCTGTCAAGGCTACACCTGCCAACCCACTTCCACAAACCCCA 660

RESULT 9
LOCUS BQ938794 975 bp mRNA linear EST 21-AUG-2002
DEFINITION AGENCOURT_8930535 NIH_MGC_94 Mus musculus cDNA clone IMAGE:6466531
ACCESSION BQ938794.1 GI:22354272
VERSION BQ938794.1 GI:22354272
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.
1 (bases 1 to 975)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: The Cepko Laboratory
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LLM13990 row: p column: 20
High quality sequence stop: 638.
Location/Qualifiers
source
1. 975
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:6466531"
/clone_lib="NIH_MGC_94"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/Note="Organ: eye; Vector: PCWV-SPORT6; Site:1; NCI;
Site:2; Salt: Cloned unidirectionally; oligo-dT primed.
Average insert size 3.3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

BASE COUNT 263 a 234 c 221 g 257 t
ORIGIN
Alignment Scores: 8.54e-75 Length: 975
Pred. No.: 672.50 Matches: 144
Score: 67.598 Conservative: 52
Percent Similarity: 49.668 Mismatches: 67
Best Local Similarity: 29.868 Indels: 27
Query Match:

```

```

DB: 14 Gaps: 8
US-10-075-846-4 (1-431) x BQ938794 (1-975)
OY 96 ValPheLeuArgGlnGlnIleTyrPheAsnProArgLeu-----SerTyrArgGlu 111
Db 1 ATTTCTTTGAGACAGAAATGAAATGAGACCCAGACTTCAAGCTTACTGACTTCAG- 57
OY 112 TyrProAspSerSerLeuAspLeuAspProSerMetLeuAspSerIleTyrPheProAsp 131
Db 58 ---CGCTACAGATGACACTGACAGAGTTCCTGTGAGCATGATGATAGTCTGTGGAACCTGAC 114
OY 132 LeuPhePheAlaAsnGlyLysGlyAlaAsnPheHisGluValThrThrAspAsnLysLeu 151
Db 115 TTATTTCTTCCAAATGAAAAAAGCAATTTTCATGATGATGACCCCAAGAAATATCTTG 174
OY 152 LeuArgIlePheLysAsnGlyAsnValLeuTyrSerIleArgLeuThrIleLeuSer 171
Db 175 TTGTTTATCTTCCGAGTGGAGAGCTCTTGTGAGCATGAGGTGTCTATTACACTTCA 234
OY 172 CysLeuMetAspLeuLysAsnPheProMetAspIleGlnThrCysThrMetGlnLeuGlu 191
Db 235 TGTCTCTAGACTTAAGTCTGTTTCCATGAGACACAGCGTGCAGAAATGCAACTTGAG 294
OY 192 SerSerSerIleLeuCysSerProLeuProSerLeuSerValGlyTyrThrMet 211
Db 295 -----AGCTTTGGATATACAAAC 312
OY 212 LysAspLeuValPheGluTyrPheGluLysPheAlaProAlaValGlnValAlaGluGlyLeu 231
Db 313 GATATTTAAGATTCATCTGCGACTGACAGAGACTCT--GTTCAAGTG--GAAAAATT 366
OY 232 ThrLeuProGlnPhe--IleLeuArgAspGlyLysAspLeuGlyCysCysThrLysHis 250
Db 367 GCTTACTCTCAATTTGATTAATAAAGAGAGATATGCAATATGAGCAACTGTACAAATATAC 426
OY 251 Tyr--AsnThrGlyLysPheThrCysIleGluValLysPheHisLeuGluArgGlnMet 269
Db 427 TATAAGGCACTGGTTACTACACTTGTGTGAGGTCTCTTCAACCCGAGAGAGACAGTT 486
OY 270 GlyTyrTyrLeuIleGlnMetTyrIleProSerLeuLeuIleValIleLeuSerTyrVal 289
Db 487 GGGTTCTACATGATGGCGGCTATATGACCAACCTTACTGATTTGGTTCTCTGCTGCTC 546
OY 290 SerPheTyrPheAsnMetAspAlaAlaProAlaArgValGlyLeuGlyIleThrVal 309
Db 547 TCTTTCTGATACACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 606
OY 310 LeuThrMetThrThrGlnSerSerGlySerArgAlaSerLeuProLysValSerTyrVal 329
Db 607 CTCAGTTTGGCTTCAGAGTGCACCCCTGCGACCGCTTCTTAAGTGTCTTATGTG 666
OY 330 LysAlaIleAspIleTyrPmeAlaValCysLeuLeuPheValPheAla--AlaLeuLeuG 349
Db 667 AAACCGCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 726
OY 349 IuTyrAlaAlaIleAsnPheValSerArgGlnHisLysGluPheIleArgLeuArgArg 369
Db 727 AGTACGCTGGGGCTCAG--GTGATGCTGAACATCCCAAAAGGTTGAAGCCGGAAG 783
OY 369 rglGlnArgArgGlnArgLeuGlu 377
Db 784 GCCCAAAATAGCTAAGGCTGAGACCA 809

RESULT 10
LOCUS AK013727 2475 bp mRNA linear HTC 19-JAN-2002
DEFINITION Mus musculus adult male hippocampus cDNA, RIKEN full-length
enriched library, clone:290059A15:gamma-aminobutyric acid (GABA-A)
receptor, subunit alpha 4, full insert sequence.
ACCESSION AK013727
VERSION AK013727.1 GI:12851204
KEYWORDS HTC; CAP trapper.

```



SOURCE Mus musculus (strain:C57BL/6J) adult male hippocampus cDNA to mRNA, clone:110-RIKEN full-length enriched mouse cDNA library

ORGANISM Mus musculus

REFERENCE 1

AUTHORS Carninci, P. and Hayashizaki, Y.

TITLE High-efficiency full-length cDNA cloning

JOURNAL Meth. Enzymol. 303, 19-44 (1999)

MEDLINE 99279253

PUBMED 10349636

REFERENCE 2

AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE 20499374

PUBMED 11042159

REFERENCE 3

AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagata, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kusunagi, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matsui, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

TITLE RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer

JOURNAL Genome Res. 10 (11), 1757-1771 (2000)

MEDLINE 20530913

PUBMED 11076861

REFERENCE 4

AUTHORS Kawai, J., Shingawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aizawa, K., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schmitt, L. M., Staudt, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamita, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombarts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyo-oka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsaki, S. and Hayashizaki, Y.

TITLE Functional annotation of a full-length mouse cDNA collection

JOURNAL Nature 409 (6821), 685-690 (2001)

MEDLINE 21085660

PUBMED 11217851

REFERENCE 5 (bases 1 to 2475)

AUTHORS Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Aizawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Horii, F., Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kono, H., Konda, M., Koya, S., Kuhlra, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Scribnal, L., Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T., Sobabe, Y., Suzuki, H., Tagami, M., Tagawa, T., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

TITLE Direct Submission

JOURNAL Submitted (10-JUL-2000) Yoshitake Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.

CDS  
gene  
CDs  
FEATURES  
source  
Location/Qualifiers  
1..2475  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="FANTOM:DB:2900059A15"  
/db\_xref="MGI:MGI:1905909"  
/db\_xref="taxon:10090"  
/clone="2900059A15"  
/sex="male"  
/tissue\_type="hippocampus"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="adult"  
1..2475  
/gene="Gabra4"  
501..2159  
/gene="Gabra4"  
/note="data source:MGI, source key:MGI:95616, gamma-aminobutyric acid (GABA-A) receptor, subunit alpha 4"  
putative"  
/codon\_start=1  
/protein\_id="PAB28975.1"  
/db\_xref="GI:12851205"  
/db\_xref="MGI:MGI:95616"  
/translation="MVSQVQVPALALCSGVSALLHFLCLACLNESPGNSKDEKLC  
PENFTRIIDSLDIDNRRLRFGCPVTEKTDYVTFSPVSDVEKEMDYDFROT  
WIDKRLTYDDEIETLRLLNMVVTVMPTDFEFGKSKSVSHMTAPKLRIRMGNTI  
LYTRRLTISACGPRRLVDPMDCACPLKGSVYPRKSEMYTGTGKPEKSEVPEKES  
SLVQYDLOITQVSESEFISITGSEYIVMIVYFHLRRKMGYFMIOFYIPCIIMVITLSQV  
SPFINKSEVPARTYFGITTVLMTTSLISARHSIPKSYATAMDPTAVCAFVFSAL  
IEFAAVYFTNIOMQAKKRIKRPPEPAVPAVLEKHTETSLONTIANLNRRTNA  
LHSESDVSKRTVEGNHNSKTSVAQESSEATPKAHLASPSPEFRANAALFMSAAAG  
LSSAASPSPGCTLRPASLGSASTRPAFGSRGRIKTYNTTGAAGNSATPPPPAPP  
SGSGTSIDKRYARLIPVTFCAFMVWVWYLSKDTMEKSSLSL"

BASE COUNT 678 a 565 c 607 g 625 t

ORIGIN

Alignment Scores:  
Pred. NO.: 9.48e-69 Length: 2475  
Score: 631.50 Matches: 151  
Percent Similarity: 51.29% Conservative: 68  
Best Local Similarity: 35.36% Mismatches: 135  
Query Match: 28.04% Indels: 73  
DB: 11 Gaps: 11

US-10-075-846-4 (1-431) x AK013727 (1-2475)

Oy 7 Alahrlleusrrprieuleuleuleutrprthleuproctlglnvalleuleuargvalala 26  
Db 555 GCCCTCGACACTCTCGTGCCTG-----CGC 581



OY	27	LeuAlaLysGIuGLuValIlysserGlyThrLysGlySerGIuPromeSerProSerasp	46
Db	582	GCTGTGTTAAACGAAATCCCCAGACAGAACTCAAAAGCAGCAAAATGTGCCCCGGAANAAT	641
OY	47	Phe-----LeuAspLysLeuMetGIyArGThrSerGIyTyrrAspAlaArgIleArg	63
Db	642	TTTAACCGGTATTCGTGACAGATTGTGCTG-----GATGGGTATGACAACAGCGTGGGT	692
OY	64	ProAsnPhelysGIyProProValAsnValIthrCysAsnIlePheIleAsnPheser	83
Db	693	CCTGATTTGGGGGCGCTGTACAGAACTGAATATATATATGTCACCGAGCTTTGGG	752
OY	84	SerValIThrLysThrMetAspTryrArgValAsnValIlePheLeuArgGIuIrrPasn	103
Db	753	CCCGTTTGTGATGTTGAATAGAAATACACTATGATGTGTCTTCGACACAGCATGATG	812
OY	104	AspProArgLeuSerTryrArgGIuIrrProAspAspSerLeuAspLeuAspProSerMet	123
Db	813	GACAAAGACATAAATAAT---GACGGCCCAATTGAANAATCTTGAGCGCTGAATATATGATG	869
OY	124	LeuAspSerIleTrypLysProAspLeuPhePheAlaAsnGIuLysGIyAlaAsnPhetis	143
Db	870	GTCACCAAAAGTTGGACCCCTGATCTTCTTCAGAAATGGAAGAAATCTGTCTCACAT	929
OY	144	GIuValIThrTrnAspAsnLysLeuLeuArgIlePheLysAsnGIyAsnValIleIrrSer	163
Db	930	AACATGACAGCTCCAAATTAAGCTTTTGAATATATGACAAATAGGACATTTTATATACAA	989
OY	164	IleArgLeuThrIleLeuLeuSerCysLeuMetAspLeuLysAsnPhProMetLaspIle	183
Db	990	ATGAGACTCACCATTAAGTCGGAGTGGCCCATGACATCGTGAGATTTCCTATGATGTT	104
OY	184	GIuIthrCysThrMetGIuLeuGIuSerSerSerIleLeuCysSerProLeuProSerIleu	203
Db	1050	CATGCGCTGCCCTTGAAATTTGGG-----	1073
OY	204	SerIleuSerValGIyTryrThrMetLysAspLeuValIleGIuIrrPheLeuGIu---	222
Db	1074	-----ACTTATGCATATCCAAAGTAGATGTCTACCTGGACCAAGGCCCTGAG	1122
OY	223	ProAlaValGIuVal---AlaGIuLysLeuThrLeuProGIuPheIleuArgAspGIu	241
Db	1128	AAGTCAGTGGAGGAGGCCAAAGAGCTCTTACGTTATGATTAATATAGACTCTCATTTGGGCAG	1187
OY	242	LysAspLeuGIyCysCysThrLysHisTryAsnThrGIyLysPheThrCysIleGIuVal	261
Db	1188	--ACTGTATCAAGGAGACATATCAAAATCTATTACAGCTGAATACATTTGATGACGGTT	1244
OY	262	LysPheHisLeuGIuArgIleMetGIyTryrTryLeuIleGIuMetTryIleProSerIleu	281
Db	1245	TACTTCCACCTCAGACGGAAGATGGGCTACTTATGATTCACAGCATATATCCATGATCATC	1304
OY	282	LeuIleValIleLeuSerTryrValSerPheTryrIleAsnMetAspAlaAlaProAlaArg	301
Db	1305	ATGACAGTGAATCTTCTTCACAGTTTCCCTTGGAATAAACAAGAGATCTGTCCACTGAGA	1364
OY	302	ValGIyLeuGIyIleThrValLeuThrMetThrIrrGIuSerSerGIySerArgAla	321
Db	1365	ACTGTATTGGAAATAACACACAGTCTCTCAGATGACACACCCTTAAGCATGACGTGGCAT	1424
OY	322	SerIleuProLysValSerTryrValLysAlaIleAspIleTryrMetAlaValCysIleuLeu	341
Db	1425	TCTTTGGCCCAAGGTCTCTATGCGACTGCCAAGGATGGTGTTCATATGCTGTGTTTGGT	1484
OY	342	PheValPheAlaAlaLeuLeuGIuITyrrAlaAlaIleAsnPhValSerArgGIuHisLys	361
Db	1485	TTTGATTTTGGCGCTTATTGAGTTTCTGCTGTCAACTAT-----	1522
OY	362	GIuPheIleArgLeuArgArgArgGIuIrrArgIrrArgIrrArgLeuGIuAspIleIleGIu	381
Db	1527	---TTACACCAACATTCAAATGACAGAAACCAAAAAGAAATATCAAAAGCT-----	1574
OY	382	GIuSerIrrGpHeTryrPheArgGIyTyrrGIyLeuGIyHisCysLeuGIuAlaIrrArgspGIy	401

[illegible]

```

Db      117 AATGAAAAAGTCCAAATTTTCATGATGACCCAGAAAAACATCCCTTTATTTT 176
Oy      156 LysAsnGlyAsnValLeuTyrSerIleArgLeuThrLeuIleLeuSerCysLeuMetAsp 175
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      177 CGTGATGGAATGCTCTTGCAGCATGAGTTATCTATTACTCTTTATGCCCCCTTGAC 236
Oy      176 LeuLysAsnPheProMetAspIleGlnThrCysThrMetGlnLeuGlnSerSerIle 195
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      237 TTGACATGTTTCCCATGATACACACGTTGCAAGATGCACATCGAG----- 284
Oy      196 LeuCysSerProLeuProSerLeuSerLeuSerValGlyTyrThrMetCysAspLeuVal 215
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      285 -----AGCTTGTGTTACACAACTGATGATTTACGA 314
Oy      216 PheGlnTyrLeuGlnAspAlaProAlaValGlnValAlaGlnGlyLeuThrLeuProGln 235
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      315 TTTATCTGCGAGTCAGAGATCC--GTGCAATTA--GAAAAAATGCTGCTGCCCA 368
Oy      236 Phe---IleLeuArgAspGlnLysAspLeuGlyCysCysThrLysHisTyr---AsnThr 253
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      369 TTGATATCAAAAAGAAAGATATTGATGATGCTAACCTGTAACAAATACTATAAGCAGC 428
Oy      254 GlyLysPheThrCysIleGlnValLysPheHisLeuGlnArgGlnMetGlyTyrTyrLeu 273
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      429 GACAAACAAAMADBKTKGTATCATCTCACCTGAGAGGAGCGCTGCGCTTTACATG 488
Oy      274 IleGlnMetTyrIleProSerLeuLeuIleValIleLeuSerTyrValSerPheThrIle 293
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      489 ATGGGGGTCTACGCCCAACCCCTGCTATGTTCTCTCTGCTGCTTCTTCTTGATC 548
Oy      294 AsnMetAspAlaAlaProAlaArgValGlyLeuGlyIleThrThrValLeuThrMetThr 313
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      549 AACCCGGACGAGTCTGCGCAGAGTCCCTGCGTATCTCTCAGTCCTGAGCTTGCC 608
Oy      314 ThrGlnSerSerGlySerArgAlaSerLeuProLysValSerTyrValLysAlaIleAsp 333
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      609 TCTGATGCAACACCTTGCGCTGAGCTTCCCAAACTTCTCTATGTAAGGCTTGAT 668
Oy      334 IleThrPheAlaValCysLeuLeuPheValPheAlaLeuLeuGlnTyrAlaAlaIle 353
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      669 GTTGGCTTATGCTTCTCTCTCTTCTTGGGTTTCTTCCCGTGAGATGCGATTGTC 728
Oy      354 -----AsnPheValSerArgGlnHisLysGlnPheIleArgLeuArgArgArg 369
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      729 CAGGTGATGCTGACACACCCCAAAAGGTTGAAGCTGAAAAAGCCAGATTCGAAGCT 788
Oy      370 GlnArg 371
      :|||||:
Db      789 GAGCAA 794

```

```

RESULT 12
LOCUS   B0042344      802 bp      mRNA      linear      EST 28-MAR-2002
DEFINITION
  UI-M-EH0-bug-p-09-0-UI.r2 NIH_BMAP_EH0 Mus musculus cDNA clone
  IMAGE:5685752 5', mRNA sequence.
ACCESSION
  B0042344
  B0042344.1 GI:19791298
VERSION  EST.
KEYWORDS
  house mouse.
SOURCE   Mus musculus.
  ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
    1 (bases 1 to 802)
  NIH-MGC http://mgc.nci.nih.gov/
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished (1999)
  Contact: Robert Strausberg, Ph.D.
  Email: cgaabs@email.nih.gov
  Tissue Procurement: Dr. James Lin, University of Iowa
  cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
  cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
  DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
  Clone Distribution: MGC clone distribution information can be

```

found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 This clone was contributed by the Brain Molecular Anatomy Project (BMAP)  
 Seq primer: pYX-5.  
 Location/Qualifiers  
 1. 802  
 /organism="Mus musculus"  
 /strain="C57BL/6"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:5685752"  
 /clone\_lib="NIH\_BMAP\_EH0"  
 /tissue\_type="whole brain"  
 /dev\_stage="embryo 18.5 dpc"  
 /lab\_host="DH10B (T1 phage resistant)"  
 /note="Organ: brain; Vector: pYX-Asc; Site: 1: EcoR I; Site: 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is CAGCCACGAC. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP). Gene discovery in the developing Mouse Nervous System, supported by National Institute of Mental Health (NIMH), Hemlin Chin, Ph.D., program coordinator."

BASE COUNT 237 a 167 c 174 g 224 t  
 ORIGIN

Alignment Scores:  
 Pred. NO.: 3.79e-68 Length: 802  
 Score: 619.50 Matches: 136  
 Percent Similarity: 62.32% Conservative: 41  
 Best Local Similarity: 47.89% Mismatches: 62  
 Query Match: 27.51% Indels: 45  
 DB: 14 Gaps: 9

```

US-10-075-846-4 (1-431) x B0042344 (1-802)
Oy      21 ValLeuLeuArgValAlaLeuAlaLysGlnGluValLysSerGlyThrLysGlySerGln 40
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      5 TTGTTGTTGAAGACCGCTTGCTTAAGAAAGTCTTCAAGAAAGGAAAGGAAAAAG 64
Oy      41 -----Promet 42
      :|||||:
Db      65 AAGCAGTACTTGTGCCATCTCAGCAGTCAACCCGAGACCTGGCCGCTGCCCCCAAC 124
Oy      43 SerProSerAspPheLeuAspLysLeuMetClyArgThrSerGlyTyrAspAlaArgIle 62
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      125 TTCACCGACATATCTTGAACAGGCTGCTG-----GTCACTTATGATCCAGCATC 175
Oy      63 ArgProAsnPheLysGlyProProValAsnValThrCysAsnIlePheIleAsnSerPhe 82
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      176 AGACCAACTTCAAGAGCAGCATCTGTTGATGATAGTACCAACATTTTATTAAATAGTTT 235
Oy      83 SerSerValThrLysThrThrMetLysPtyrArgValAsnValPheLeuArgGlnGlnTyr 102
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      236 GATCCATTTCAGAGACACAACTAGACTTAACTTTCTTGACAGACAGAAATGG 295
Oy      103 AsnAspProArgLeu-----SerTyrArgGlnTyrProAspAspSerLeuAsp 118
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      296 AATGACCCACAGACTCAAGTACCTAGTGCATTCAGA-----GGCTCAGATGCACCTACA 349
Oy      119 LeuAspProSerMetLeuAspSerIleTyrPtyrProAspLeuPhePheAlaAsnGlnLys 138
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      350 GTTACCCACACATGTAAGTCTTGTGGAACACTTATCTTTGCAAAAGAAAAA 409
Oy      139 GlyAlaAsnPheHisGlnValThrThrAspAsnLysLeuLeuArgIlePheLysAsnGly 158
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

Db		410	AGTGCACAAATTTTCATGANGTGACCCAGAANAATAATTCCTGTATTACTTTCGGGGATCGA	465
Oy		159	AsnValIeuTYrSerLleArgLeuThrLleuLleuSerCysLeuMetAspLeuLysasn	178
Db		470	GAGCTCTCTTGAGACATAGAGTTCTCATTAATCACTTTCATCTCCCTAGACTAACCTCTG	529
Oy		179	PheProMetAspRIleGlnThrCysThrMetInLenuGlusSerSerLileucCysSer	198
Db		530	TTCGCCAGTGACACACAACAACGCTGCAGAAATGCAACTTAG-----	568
Oy		199	ProIeuProSerLeuSerLieuSerValGIlyTrThrMetLysAspLeuValPheGluTrp	218
Db		569	-----AGCTTTGCGAATATCAACACCAGATGAATTAAGATTATCATCTGG	607
Oy		219	LengLuAspAlaProAlaValGlnValAlaGluLyuLyuTruLnuProgInPhe----lle	237
Db		608	CAGTCAGACAGATCTCT---GTTCACTTG---GAAAAAATGCTTACTTACCTCAATTTGATNT	661
Oy		238	LeuArGaspGluLysAspLeuGluYCysScsThryshstYr---AsnThrGlyLysPhe	256
Db		662	AAAAGAGGAGATATCGAATATATGCAACGTCACAAATAATCTATAAGCAGACGGTTACTAC	721
Oy		257	ThrcyslIlegluValLysPheHisLeuGluLnuArgInmetGIlyTrYLleuIlleGlnMet	276
Db		722	ACTGTGTGGAGGTCATCTTACCCCTGAGGACAGGTTGGTGGTTCTACATGATGGCGTA	781
Oy		277	TyrIleProser	280
Db		782	TATGCACCACT	793
RESULT_13				
AUI33223				
LOCUS		817 bp	mRNA	linear EST 01-AUG-2002
DEFINITION		AUI33223 NT2RP4 Homo sapiens cDNA clone NT2RP4001569 5', mRNA		sequence.
ACCESSION		AUI33223		
VERSION		AUI33223.1	GI:10993762	
KEYWORDS		EST.		
SOURCE		human.		
ORGANISM		Homo sapiens		
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
		Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
REFERENCE		1 (bases 1 to 817)		
AUTHORS		Ota,T., Sugiyama,T., Ishii,S., Suzuki,Y., Saico,K., Yamamoto,J., Nishikawa,T., Nakamura,Y., Nagai,T., Sugano,S., Masuno,Y. and Isoga,T.		
TITLE		HRI human cDNA project (Ota,T., Sugiyama,T., Ishii,S., Suzuki,Y., Saico,K., Yamamoto,J., Nishikawa,T., Nakamura,Y., Nagai,T., Sugano,S., Masuno,Y., Isoga,T.)		
JOURNAL		Unpublished (2000)		
COMMENT		Contact: Takao Isoga Genomics Laboratory Helix Research Institute 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan Tel.: 81-438-52-3975 Fax: 81-438-52-3986 Email: genomics@hri.co.jp HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.  Location/Qualifiers 1..817 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="NT2RP4001569" /clone_id="NT2RP4" /cell_type="teratocarcinoma" /cell_line="NT2" /note="vector: pME18SFJ3; mRNA from NT2 neuronal precursor cells after 2-weeks retinoic acid (RA) induction"		
BASE COUNT		207 a	207 c	198 g 202 t 3 others
ORIGIN				

Alignment Scores:		4,86e-67	Length:	817
Pred. No.:	Score:	611.00	Matches:	121
Percent Similarity:	62.14%	Conservative:	53	
Best Local Similarity:	43.21%	Mismatches:	90	
Query Match:	27.13%	Indels:	16	
DB:	9	Gaps:	3	
US-10-075-846-4 (1-431) x AU133223 (1-817)				
OY	57	GLYTRASPAALaRgIIleIaRgPRoKsnPhelYsgIyPRoPovaIAsnValIhCysAsn	76	
Db	15	GGCTTACGACATTTGGCCTTAAGACCAGACTTCCGGGGTCCCGGTCTGGGTGGGATGAC	74	
OY	77	IIepheIIeaSnSerPheSerSerValIhTrIystrIhTmLeaSPtYrArgValAsnVal	96	
Db	75	ATGCACATCGCAGCATGCACATGGTTTCCGAAGTCAACATGATATTACCTTAACATG	134	
OY	97	PhelAuAaRgIIngIInTPraSnAsPProARgIeUseryTyrARgIuTyPRoAspAsPser	116	
Db	135	TATTTCACACATATTGGAGATTAAGAGCTCCGCTTTCTGGATCCCT--CTCAAC	191	
OY	117	LeuAsPLeaAsPProSerMetLeuAsPserIIerTPlYsPRoAsPLeuPhehAlaSn	136	
Db	192	CTCACGCTTGAACAACCGAGCTGGCTGTACCAGCTATGGTGGCCGACACATATTTCTTAAT	251	
OY	137	GIuTySGIyIaAsnPhenIISgluValIhTrIhAsPaSnLYsLeuIeuaRgIIepheIys	156	
Db	252	GACAAAAAGTATTTGTGCATGAGATGACAGTGAAMAACCCATGATCCGCTTCACCT	311	
OY	157	AsnGIyAsnValIeUtySerSerIIeRgIeUtyrIleUtyIleUserCysIeUmetrIleu	176	
Db	312	GATGGAGACAGTGCCTGTATGGGCTCGAGATTCACACAGACAGCATGATGATGAGACTTC	371	
OY	177	LYsAsnPhEPrometaSPleGIInThrcYstrIhMetGIInleuGIuSerSerIleUeu	196	
Db	372	AGGAGATACCCCTGGAGCAGACAGCACTCGTGAATTGA-----	416	
OY	197	CysSerPProAsPProSerLeuSerLeuSerValGIyTrIhTmLeTyAsPLeuValPhe	216	
Db	417	-----AGCTATGGCTACACACAGATGACATTGAGTTT	449	
OY	217	GIuTrPLeuGIuAsPAlaProAlaValGIInValaIaGIuGIyLeuThIleuProGIInPhe	236	
Db	450	TACTGGCGAGCGGGACAGAGCTGTTCACGGAGGAAAGATGATGACTCCCGCAGTTC	509	
OY	237	IIleuAuAgaRgIuLYsAsPLeuIyCysCysThIyShIStYrAsnThIyGIyAsPhe	256	
Db	510	TCATTCGTGGAGCACCGT--CTGCTCTCGAGGAATGTTGCTTCCGCCACAGTCCCTAT	566	
OY	257	ThrcYsIIeGIuValLYsPheIISleuGIuARgIeUmetGIyTrIyTrIleuIleGIInmet	276	
Db	567	CCGAGACTGTACAGAGCTTTCGGTGTAAAGAGAACATTGGATCTTCATTTCCAGACT	626	
OY	277	TyrIleProSerLeuIleValIleUserTrIpaIserPheTrIleAsnMetLsp	296	
Db	637	TATATGCCCTTATACGATGAACGTTTGTCTGGGGGTCCCTTCGATCAATTATCAT	686	
OY	297	AlaIaIaPProIaAaRgValGIyLeuGIyIleThrTrIhValIleUthMetThThIcIInSer	316	
Db	687	GCATCTGCTGTAGAGTTCGCCCTCGGATCAACAACATGGGTGACATACACCACTTNAAC	746	
OY	317	SerGIySerAaRgAlaSerLeuProLYsValSerTyRValLYsAlaIlaSPleRmet	336	
Db	747	ACCCACACCTTCGAGAGACCTTCCCAAAATNCCCTATGTCAAAAGCCATTGACATGACTT	806	
RESULT 14				
LOCUS	BC029850	1942 bp	mRNA	linear
DEFINITION	Homo sapiens, gamma-aminobutyric acid (GABA) A receptor, gamma 3, clone IMAGE:5178575, mRNA.			
ACCESSION	BC029850			



AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT Genoscope - Centre National de Sequencage  
 BP 191 91006 Evry cedex - France  
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr.

FEATURES  
 source 1. 948  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="CS0DF03YM17"  
 /clone\_1ib="LRI\_Fl013\_FBRn1"  
 /dev\_stage="pooled tissue from post conception fetuses (20 week, 24 week and 26 week)"  
 /lab\_host="DH10B"  
 /note="Organ: Fetal brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-OLIGO(4T) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was constructed by life technologies. Contact: Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax: (1) 301 610 8371 Email: fliang@life.com URL: http://fulllength.invitrogen.com"

BASE COUNT 275 a 215 c 210 g 247 t 1 others  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1.03e-61 Length: 948  
 Score: 570.50 Matches: 121  
 Percent Similarity: 57.64% Conservative: 60  
 Best Local Similarity: 38.54% Mismatches: 108  
 Query Match: 25.33% Indels: 25  
 DB: 9 Gaps: 6

US-10-075-846-4 (1-431) x AL539373 (1-948)

```

Oy 10 SerPheLeuLeuLeuThrPheProGlyGlnVal-----LeuLeuArgValAla 26
    ||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 63 TCAGAGATTAGTATGTCGACAGTACAAAATCGAGAGAGCTGCGGCTTCTCTTCCT 122
Oy 27 LeuAlaLysGlnGluValLysSerGlyThrLysGlySerGlnPheMetSerProser--- 45
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 123 GTGATGATTACCATGCTGTGTGTCACACACACACCAATGAACCCAGCAACATGTCATAC 182
Oy 46 -----AspPheLeuAspLysLeuMetGlyArgThrSerGlyTyrAspAlaArgIleArg 63
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 183 GTGAAGAGACAGTCGACAGATTGCTC-----AAAGATATGACATTGCTTCGCG 233
Oy 64 ProAsnPhelYsGlyProProValAsnValThrCysAsnIlePheIleAsnSerPheSer 83
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 234 CCGGACTTCGGAGGCCCCCGCTCGACGCTGGGATCGGATGATGTCGCCAGCATAGAC 293
Oy 84 SerValThrLysThrThrMetAspTyrArgValAsnValPheLeuArgGlnGlnTyrAsn 103
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 294 ATGCTCTCCGAGTGAATGATGATATACATCACCATGTATTTCCAGCAGCTTGGA 353
Oy 104 AspProAlaGluSerTyrArgIleUtyrProAspAspSerLeuAspLeuAspProSerMet 123
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 354 GACAAAGAGCTTCTTATCTGAAATCCCA--CTGAACCTCACCCCTAGACAAATAGGCTA 410
Oy 124 LeuAspSerIleTyrLysProAspLeuPhePheAlaAsnGlnLysGlyValAsnPheHis 143
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 411 GCTGACCAACTCTCGGTACCAAGACACTTCTTGAATGACAGAAATCATTTGTGCAT 470
Oy 144 GluValThrThrAspAsnLysLeuLeuArgIlePheLysAsnGlyValAsnValLeuTyrSer 163
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 471 GGGGTACACGTAAGAAATGATGATGTCATCCGATGATGGAACAGTTCTCTATGGA 530
Oy 164 IleArgLeuThrLeuIleLeuSerCysLeuMetAspLeuLysAsnPheProMetAspIle 183
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

```

Db 531 CTCGCAATCACAACACAGCTCATGTATGATGATCTTCGAAGATATTCACATGATGAG 590
Oy 184 GlnThrCysThrMetGlnLeuGlnUserSerSerIleLeuCysSerProLeuProSerLeu 203
    ||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 591 CAGAACTGCACCCCTGAGATCGAA-----
Oy 204 SerLeuSerValGlyTyrThrMetLysAspLeuValPheGluTyrPheGlnAspAlaPro 223
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 615 -----AGTTATGGCTATPCCACTGATGACATTTTACTGGAATGAGAGACAGG 668
Oy 224 AlaValGlnValAlaGlnGlyLeuThrLeuProGlnPheIleLeuArgAspGlnLysAsp 243
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 669 GCAGTACACTGCTGTAATTAATTCGAACTCTCTCAATTTTCAATTGTGACAAAGATG 728
Oy 244 LeuGlyCysCysThrLysHisTyrAsnThrGlyLysPheThrCysIleGluValLysPhe 263
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 729 GTGCTTAAGAGGTGAG--TTCAACACAGAGGCTATCCAGCAGCTGCATTAAGTTT 785
Oy 264 HisLeuGluArgGlnMetGlyTyrThrLeuIleGlnMetTyrIleProSerLeuLeuIle 283
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 786 CGCTTAAGAGAAACATTGCTTACTTCATTTTGCAAACCTACATGCCCTTACACTGATT 845
Oy 284 ValIleLeuSerTyrValSerPheThrPheAsnMetAspAlaAlaProAlaArgValGly 303
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 846 ACAATTCTGTCTCTGCGGTGCTTTTGATCAACATGATGATCTGCAAGCCAGACTCCCA 905
Oy 304 LeuGlyIleThrThrValLeuThrMetPheThrGlnSerSer 317
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 906 CTAGGATTCGACACATCTTTTACAATGACACACATCAGCAC 947

```

Search completed: July 1, 2003, 00:55:27  
 Job time : 2277.88 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: June 30, 2003, 22:23:50 ; Search time 131.866 Seconds  
(without alignments)  
3531.204 Million cell updates/sec

Title: US-10-075-846-4\_COPY\_192\_207  
Perfect score: 75  
Sequence: 1 SSSILCSPLPSLSLV 16

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlp  
-O=/cgn2.1/USPRO\_spool/US10075846/runat\_25062003.163648.5147/app\_query.fasta\_1.782  
-DB=genembl -QFRT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blousum62 -TRANS=human40.cdi -List=45  
-DOCLALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10075846 -CGCN\_1.1.3854 -runat.25062003.163648.5147 -ICPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Genembl:.\*  
1: gb.ba:.\*  
2: gb.htg:.\*  
3: gb.in:.\*  
4: gb.om:.\*  
5: gb.ov:.\*  
6: gb.pat:.\*  
7: gb.ph:.\*  
8: gb.pl:.\*  
9: gb.pr:.\*  
10: gb.ro:.\*  
11: gb.sts:.\*  
12: gb.sy:.\*  
13: gb.un:.\*  
14: gb.vl:.\*  
15: em.ba:.\*  
16: em.fun:.\*  
17: em.hum:.\*  
18: em.in:.\*  
19: em.mu:.\*  
20: em.om:.\*  
21: em.or:.\*  
22: em.ov:.\*  
23: em.pat:.\*  
24: em.ph:.\*  
25: em.pl:.\*  
26: em.ro:.\*  
27: em.sts:.\*  
28: em.un:.\*

29: em.vl:.\*  
30: em.htg\_hum:.\*  
31: em.htg\_inv:.\*  
32: em.htg\_other:.\*  
33: em.htg\_mus:.\*  
34: em.htg\_pln:.\*  
35: em.htg\_fod:.\*  
36: em.htg\_mam:.\*  
37: em.htg\_vrt:.\*  
38: em.sy:.\*  
39: em.htgo\_hum:.\*  
40: em.htgo\_mus:.\*  
41: em.htgo\_other:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	68	90.7	39796	6 AX037570	AX037570 Sequence
2	68	90.7	39796	9 HSU3563	293848 Human DNA s
3	61	81.3	170676	9 AC026258	AC026258 Homo sapi
4	61	81.3	183125	9 AP002848	AP002848 Homo sapi
5	61	81.3	189127	9 AC084114	AC084114 Homo sapi
6	61	81.3	195477	2 AC079196	AC079196 Homo sapi
7	60	80.0	91582	2 AC109887	AC109887 Rattus no
8	57	76.0	169845	2 AC095732	AC095732 Rattus no
9	57	76.0	191204	2 AC094926	AC094926 Rattus no
10	56	74.7	104525	2 AC095220	AC095220 Rattus no
11	56	74.7	172555	2 AC102866	AC102866 Mus muscu
12	55	72.3	598	5 G6126MRP	X61200.G.gallus 12
13	55	72.3	139754	2 AC118327	AC118327 Rattus no
14	55	72.3	164772	2 AC095184	AC095184 Rattus no
15	55	72.3	174354	2 AC108830	AC108830 Mus muscu
16	55	72.3	193627	2 AC091427	AC091427 Mus muscu
17	55	72.3	197676	2 AC114635	AC114635 Mus muscu
18	55	72.3	217164	2 AC094182	AC094182 Rattus no
19	55	72.3	238033	2 AC125309	AC125309 Mus muscu
20	55	72.3	254752	2 AC127336	AC127336 Mus muscu
21	55	72.3	310148	2 AC073729	AC073729 Mus muscu
22	54	72.0	309	9 HS255ZEP	Z17151.H.sapiens
23	54	72.0	14839	1 AE000738	AE000738 Aquifex a
24	54	72.0	69740	2 AC087748	AC087748 Homo sapi
25	54	72.0	85149	9 AL137879	AL137879 Human DNA
26	54	72.0	106186	2 AC105876	AC105876 Rattus no
27	54	72.0	113704	9 HS267P19	275889 Human DNA s
28	54	72.0	117628	9 AC008891	AC008891 Homo sapi
29	54	72.0	120652	2 AL353724	AL353724 Human DNA
30	54	72.0	121705	2 AC118360	AC118360 Rattus no
31	54	72.0	153042	9 AP002984	AP002984 Homo sapi
32	54	72.0	153788	2 AC118805	AC118805 Rattus no
33	54	72.0	164472	9 AC018897	AC018897 Homo sapi
34	54	72.0	178733	2 AC068786	AC068786 Homo sapi
35	54	72.0	179217	2 AC099748	AC099748 Bos tauru
36	54	72.0	180226	2 AC109001	AC109001 Rattus no
37	54	72.0	209621	2 AC111125	AC111125 Mus muscu
38	54	72.0	223150	2 AC104396	AC104396 Mus muscu
39	54	72.0	233454	2 AC096987	AC096987 Rattus no
40	54	72.0	263744	2 AF401201	AF401201 Mus muscu
41	53.5	71.3	35412	8 SPAC4F8	298530 S.pombe chr
42	53.5	71.3	174005	2 AC117603	AC117603 Mus muscu
43	53	70.7	549	4 AB048310	AB048310 Equus cab
44	53	70.7	3301	6 AX333072	AX333072 Sequence
45	53	70.7	3301	9 HUWMPARA	L07592 Human perox

RESULT 1

ALIGNMENTS

AX037570  
LOCUS AX037570 39796 bp DNA linear PAT 16-NOV-2000  
DEFINITION Sequence 6 from Patent WO0058461.  
ACCESSION AX037570  
VERSION AX037570.1 GI:11226991  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE  
AUTHORS Rappold-Hoerbrand,G.  
TITLE Gene for ataxia  
JOURNAL Patent: WO 0058461-A 6 05-OCT-2000;  
RAPPOLD HOERBRAND GUERIN (DE)  
FEATURES  
source 1.39796  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 10630 a 9337 c 9355 g 10474 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 9.92 Length: 39796  
Score: 68.00 Matches: 14  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 93.33% Mismatches: 0  
Query Match: 90.67% Indels: 0  
DB: 6 Gaps: 0  
US-10-075-846-4\_COPY\_192\_207 (1-16) x AX037570 (1-39796)  
QY 2 SerSerIleLeuCySerProLeuProSerLeuSerLeuSerVal 16  
Db 35719 GCATCATACTCTCGACGCCCTCGCATCTCTGTCACTTCACTT 35763  
RESULT 2  
HSU3563/c 39796 bp DNA linear PRI 23-NOV-1999  
LOCUS HSU3563/c  
DEFINITION Human DNA sequence from cosmid U3563, between markers DXS366 and  
ACCESSION DXS87 on chromosome X contains ESTs.  
VERSION 293848  
KEYWORDS 293848.1 GI:1934909  
X.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE  
AUTHORS Burgess,J.  
TITLE Direct Submission  
JOURNAL Submitted (07-APR-1997) Sanger Centre, Hinxton, Cambridgeshire,  
CB10 1RQ, UK. E-mail enquiries: humquery@sanger.ac.uk  
COMMENT  
requests: clonerequest@sanger.ac.uk  
IMPORTANT: This sequence is the entire insert of clone U3563. This  
sequence has been finished according to sequence map criteria as  
follows. An attempt is made to resolve all sequencing problems,  
such as compressions and repeats, but not necessarily within known  
annotated human repeat sequence elements (e.g. Alu). Where the  
sequence is ambiguous, there is an annotation using the 'unsure'  
feature key.  
The true left end of clone U3563 is at 1 in this sequence. The true  
right end of clone U3563 is at 39796.  
U3563 is from the Lawrence Livermore National Laboratory  
flow-sorted X chromosome cosmid library LLOXNC01.  
Location/Qualifiers  
1.39796  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="X"  
/map="X"  
/clone="LLOXNC01-3563"  
/clone\_lib="LLOXNC01"

repeat\_region 9..107  
repeat\_region /note="MIR2 repeat: matches 48..146 of consensus"  
1267..1477  
/note="MIR repeat: matches 1..218 of consensus"  
2267..2446  
repeat\_region /note="MIR repeat: matches 196..19 of consensus"  
2619..2734  
repeat\_region /note="L1PA6 repeat: matches 571..457 of consensus"  
2863..3051  
repeat\_region /note="MER5A repeat: matches 7..189 of consensus"  
3608..3745  
repeat\_region /note="MIR repeat: matches 20..156 of consensus"  
6949..7115  
repeat\_region /note="MIR repeat: matches 212..46 of consensus"  
7944..8166  
repeat\_region /note="MIR repeat: matches 224..2 of consensus"  
8750..8823  
repeat\_region /note="MIR repeat: matches 65..138 of consensus"  
14745..14874  
repeat\_region /note="MIR2 repeat: matches 146..21 of consensus"  
15339..15398  
repeat\_region /note="5 copies of 12 mer 82 & conserved"  
16847..17148  
repeat\_region /note="AluSg repeat: matches 1..302 of consensus"  
17947..18235  
repeat\_region /note="AluYb repeat: matches 295..1 of consensus"  
20312..20454  
repeat\_region /note="MIR repeat: matches 20..162 of consensus"  
20456..20696  
repeat\_region /note="AluSx repeat: matches 44..284 of consensus;  
Incomplete repeat"  
21916..22233  
repeat\_region /note="AluSp repeat: matches 1..302 of consensus"  
22428..22729  
repeat\_region /note="AluSx repeat: matches 1..302 of consensus"  
22816..22915  
repeat\_region /note="MER5A repeat: matches 71..169 of consensus"  
23170..23463  
repeat\_region /note="AluSx repeat: matches 300..2 of consensus"  
23578..23716  
repeat\_region /note="MIR2 repeat: matches 16..146 of consensus"  
23844..23993  
repeat\_region /note="MIR2A1 repeat: matches 365..215 of consensus"  
24402..24703  
repeat\_region /note="AluYb repeat: matches 302..11 of consensus"  
24710..24794  
repeat\_region /note="MIR repeat: matches 110..20 of consensus"  
25054..25347  
repeat\_region /note="AluSg repeat: matches 299..1 of consensus"  
25430..25730  
repeat\_region /note="AluSp repeat: matches 303..1 of consensus"  
25733..26033  
repeat\_region /note="AluSx repeat: matches 302..2 of consensus"  
26644..26753  
repeat\_region /note="FLAM\_A repeat: matches 123..10 of consensus"  
26869..27014  
repeat\_region /note="MIR repeat: matches 49..196 of consensus"  
27270..27486  
repeat\_region /note="MIR repeat: matches 31..250 of consensus"  
27575..27798  
repeat\_region /note="AluYb repeat: matches 87..302 of consensus;  
Incomplete repeat"  
30559..30679  
repeat\_region /note="AluSg repeat: matches 1..121 of consensus;  
Incomplete repeat"  
30679..30799  
repeat\_region /note="AluSg repeat: matches 124..1 of consensus;  
Incomplete repeat"  
30706..30781  
repeat\_region /note="L1MB8 repeat: matches 413..676 of consensus"  
31110..31403  
repeat\_region /note="AluSg repeat: matches 1..293 of consensus"



```

repeat_region 31775..32116
/note="MIR181 repeat: matches 365. .1 of consensus"
repeat_region 32540..32840
/note="AluSP repeat: matches 1. .302 of consensus"
repeat_region 34108..34350
/note="MIR repeat: matches 247. .6 of consensus"
repeat_region 34410..34709
/note="AluSP repeat: matches 301. .1 of consensus"
repeat_region 34732..34831
/note="MIR repeat: matches 106. .203 of consensus"
repeat_region 34850..34976
/note="MIR repeat: matches 36. .122 of consensus"
repeat_region 36739..36891
/note="AluSP repeat: matches 148. .300 of consensus"
repeat_region 36958..37110
/note="MIR repeat: matches 105. .260 of consensus"
repeat_region 37468..38250
/note="MER21B repeat: matches 1. .770 of consensus"
prim_transcript <38771..>39796
/note="match: 5' EST C16803 clone GEN-530B01: match: 5'
EST C16720 clone GEN-522A09: match: 5' EST C17570 clone
GEN-549E01"
prim_transcript <38771..>39418
/note="match: 5' EST T86096 clone 114200: match: 3' EST
M67655 clone 343194"
BASE COUNT 10474 a 9355 c 9337 g 10630 t
ORIGIN
Alignment Scores:
Pred. No.: 9.92 Length: 39796
Score: 68.00 Matches: 14
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 93.33% Mismatches: 0
Query Match: 90.67% Indels: 0
DB: 9 Gaps: 0
US-10-075-846-4_COPY_192_207 (1-16) x HS05G3 (1-39796)
Oy 2 SerSerIleuCySerProleuProSerIleuSerIleuSerVal 16
Db 4078 GCATTCATCTCTGCGAGCCCTCTGCGATCTGTGACCTTTCAGTT 4034
RESULT 3
AC026258 170676 bp DNA linear PRI 18-AUG-2001
LOCUS Homo sapiens chromosome 8, clone RP11-67N21, complete sequence.
DEFINITION AC026258
AC026258.4 GI:14579802
VERSION AC026258.4 GI:14579802
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 170676)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Unpublished
JOURNAL 2 (bases 1 to 170676)
AUTHORS Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavsky,L., Boukhalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Coyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kam,L., Karatas,A.,
Klein,J., Laroque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meldrim,J., Menes,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,

```

```

O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tessie,S., Theodore,J., Tirrell,A., Travers,M., Triggillo,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
JOURNAL Submitted (21-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 170676)
REFERENCE
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., All,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,P.,
Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P.,
Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gort,S.,
Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A.,
Horton,L., Hulme,M., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
Karatas,A., Kells,C., Laroque,K., Lamazares,R., Landers,T.,
Lehoczy,J., Levine,R., Lieu,G., Maclean,C., Macdonald,P., Major,J.,
Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,
McPheeters,R., Meldrim,J., Menes,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nordu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Reta,R.,
Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,
Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S., Severy,P.,
Sougnuez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tessie,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Triggillo,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
JOURNAL Submitted (18-AUG-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 1, 2001 this sequence version replaced by:14336597.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L7482
Center clone name: 67_N21
-----
FEATURES
SOURCE
Location/Qualifiers
1..170676
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8"
/clone="RP11-67N21"
/clone_lib="RPCI-11 Human Male BAC"
79..293
/rpt_family="MERS8"
complement(913..1811)
/rpt_family="L1ME2"
1929..2057
/rpt_family="L1ME2"
2238..2911
/rpt_family="L1ME2"
2916..2965
/rpt_family="AT-rich"
3230..3303
/rpt_family="L2"
/rpt_family="L2"
complement(3538..3699)
/rpt_family="MLT101"
complement(3924..4315)
/rpt_family="MSTB"

```

```

repeat_region 4623..4950 /rpt_family="L1MED"
repeat_region 4951..5281 /rpt_family="L1MED"
repeat_region 5282..5492 /rpt_family="Aluv"
repeat_region 5494..5606 /rpt_family="L1MED"
repeat_region 5607..6129 /rpt_family="L1MED"
repeat_region 6130..6490 /rpt_family="L1MED"
repeat_region 6491..6568 /rpt_family="THE1B"
repeat_region 6562..6859 /rpt_family="L1ME1"
repeat_region 6863..7181 /rpt_family="MERS8A"
repeat_region 7569..7706 /rpt_family="L1MED"
repeat_region 7966..8197 /rpt_family="L1MED"
repeat_region 8563..8584 /rpt_family="L1MED"
repeat_region 8672..8718 /rpt_family="AT_rich"
repeat_region 8807..8834 /rpt_family="(TATG)n"
repeat_region /rpt_family="AT_rich" complement(9885..10092)
repeat_region /rpt_family="L2" complement(10436..10495)
repeat_region /rpt_family="L2" 10763..10789
repeat_region /rpt_family="AT_rich" complement(10813..10901)
repeat_region /rpt_family="L2" 11227..11953
repeat_region /rpt_family="L2" 11954..12264
repeat_region /rpt_family="Alusx" 1265..12501
repeat_region /rpt_family="L2" complement(12500..12569)
repeat_region /rpt_family="MIR" 13809..13858
repeat_region /rpt_family="(TG)n" 14086..14159
repeat_region /rpt_family="MIR" 15633..15730
repeat_region /rpt_family="L2" complement(18562..18744)
repeat_region /rpt_family="MIR" 18878..19453
repeat_region /rpt_family="MULTIF" 19809..19825
repeat_region /rpt_family="MERS6" complement(19826..20188)
repeat_region /rpt_family="THE1B" 20189..20443
repeat_region /rpt_family="MERS6" 20447..20480
repeat_region /rpt_family="(TTG)n" 20504..20636
repeat_region /rpt_family="MERS6" 22697..22726
repeat_region /rpt_family="(TG)n" 23655..23830
repeat_region /rpt_family="MIR" 25146..25267
repeat_region /rpt_family="(TA)n" 25472..25500
repeat_region /rpt_family="AT_rich" 27066..27289

```

```

repeat_region /rpt_family="MIR" complement(27340..27437)
repeat_region /rpt_family="L2" complement(27520..27655)
repeat_region /rpt_family="MIR" 27834..28144
repeat_region /rpt_family="AlusC" 28459..28616
repeat_region /rpt_family="L2" complement(28851..29045)
repeat_region /rpt_family="MERS20" complement(29201..29416)
repeat_region /rpt_family="L2" 30415..30513
repeat_region /rpt_family="GA_rich" complement(30784..30992)
repeat_region /rpt_family="MIR3" complement(31406..31478)
repeat_region /rpt_family="MERS16" 33152..33240
repeat_region /rpt_family="L2" complement(33241..33407)
repeat_region /rpt_family="MERSA" 33408..33810
repeat_region /rpt_family="L2" complement(33832..34249)

```

```

Alignment Scores:
Pred. No.: 506 Length: 170676
Score: 61.00 Matches: 12
Percent Similarity: 93.33% Conservative: 2
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 81.33% Indels: 0
DB: 9 Gaps: 0

```

US-10-075-846-4\_COPY\_192\_207 (1-16) x AC026258 (1-170676)

```

Oy 1 SerSerIleLeuCySerProleuProSerIleuSer 15
::: |||||
Db 46678 ACACATCATCTCTGCTCACCCTGCCCAACCTCACCTCTCT 46722

```

```

RESULT 4
AP002848 183125 bp DNA linear PRI 20-JAN-2001
LOCUS Homo sapiens genomic DNA, chromosome 8q23, clone:KB1670H7.
DEFINITION AP002848
ACCESSION AP002848
VERSION AP002848.2 GI:12328451
KEYWORDS
SOURCE Homo sapiens cell_line:FLEB 14 - 14 DNA, clone_id:Keio BAC library
clone:KB1670H7.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 183125)
AUTHORS Shimizu,N. and Asakawa,S.
TITLE Homo sapiens chromosome clone KB1670H7 on 8q23
JOURNAL Published Only in Database (2000) In press
REFERENCE 2 (bases 1 to 183125)
AUTHORS Shimizu,N. and Asakawa,S.
TITLE Direct Submission
JOURNAL Submitted (10-AUG-2000) Nobuyoshi Shimizu, Keio University, School
of Medicine, Molecular Biology, 35 Shinanomachi, Shinjuku-ku, Tokyo
160-8582, Japan (E-mail:nshimizu@mb.med.keio.ac.jp,
Tel:81-3-3351-2370, Fax:81-3-3351-2370)
On Jan 19, 2001 this sequence version replaced gi:9798632.

```

```

COMMENT
FEATURES
Location/Qualifiers
1..183125
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8q23"
/clone="KB1670H7"
/cell_line="FLEB 14 - 14"

```

repeat_region	/clone_lib="Kelo BAC library"	repeat_region	complement(20412..21003)
	complement(119..1004)		/evidence-not_experimental
	/rpt_family="L1PA13"	repeat_region	/rpt_family="HERV16"
	complement(5651..5736)		complement(21009..21269)
	/evidence-not_experimental		/evidence-not_experimental
repeat_region	/rpt_family="L2"	repeat_region	/rpt_family="L1PA16"
	complement(5870..7113)		complement(21559..21801)
	/evidence-not_experimental		/evidence-not_experimental
repeat_region	/rpt_family="HERVL"	repeat_region	/rpt_family="NAL1"
	complement(7215..8911)		22760..22790
	/evidence-not_experimental		/evidence-not_experimental
repeat_region	/rpt_family="HERVL"	repeat_region	/rpt_family="(TC)n"
	complement(9824..9923)		22791..22833
	/evidence-not_experimental		/evidence-not_experimental
repeat_region	/rpt_family="MIR"	repeat_region	/rpt_family="(CA)n"
	10182..10506		22853..22873
	/evidence-not_experimental		/evidence-not_experimental
repeat_region	/rpt_family="L2"	repeat_region	/rpt_family="AT_rich"
	10507..10798		23766..23825
	/evidence-not_experimental		/evidence-not_experimental
repeat_region	/rpt_family="AlusX"	repeat_region	/rpt_family="L2"
	10799..10895		complement(26443..26521)
	/evidence-not_experimental		/evidence-not_experimental
repeat_region	/rpt_family="L2"	repeat_region	/rpt_family="MADE1"
	11776..11956		26702..26766
	/evidence-not_experimental		/evidence-not_experimental
repeat_region	/rpt_family="MIR"	repeat_region	/rpt_family="MLT1J"
	13212..13232		26767..27179
	/evidence-not_experimental		/evidence-not_experimental
repeat_region	/rpt_family="(TC)n"	repeat_region	/rpt_family="MLT2G"
	complement(13752..13836)		27180..27403
	/evidence-not_experimental		/evidence-not_experimental
repeat_region	/rpt_family="MIR"	repeat_region	/rpt_family="MLT1J"
	complement(13969..14098)		27419..27524
	/evidence-not_experimental		/evidence-not_experimental
repeat_region	/rpt_family="L2"	repeat_region	/rpt_family="L2"
	complement(14307..14557)		complement(27528..27609)
	/evidence-not_experimental		/evidence-not_experimental
repeat_region	/rpt_family="MIR"	repeat_region	/rpt_family="L2"
	complement(14579..14677)		27830..27862
	/evidence-not_experimental		/evidence-not_experimental
repeat_region	/rpt_family="L2"	repeat_region	/rpt_family="AT_rich"
	14820..14998		complement(28038..28300)
	/evidence-not_experimental		/evidence-not_experimental
repeat_region	/rpt_family="MLT1H"	repeat_region	/rpt_family="MLT1K"
	complement(15176..15653)		complement(28397..28851)
	/evidence-not_experimental		/evidence-not_experimental
repeat_region	/rpt_family="L1PA2"	repeat_region	/rpt_family="MER70A"
	complement(15666..15940)		29183..29494
	/evidence-not_experimental		/evidence-not_experimental
repeat_region	/rpt_family="MLT1B"	repeat_region	/rpt_family="AluY"
	complement(15994..16116)		29731..29753
	/evidence-not_experimental		/evidence-not_experimental
repeat_region	/rpt_family="MLT1B"	repeat_region	/rpt_family="(TTTTC)n"
	16746..16902		complement(30040..30244)
	/evidence-not_experimental		/evidence-not_experimental
repeat_region	/rpt_family="MER5A"	repeat_region	/rpt_family="MIR"
	17054..17186		complement(30280..30581)
	/evidence-not_experimental		/evidence-not_experimental
repeat_region	/rpt_family="MLT2E"	repeat_region	/rpt_family="AluJb"
	17187..17595		30861..30888
	/evidence-not_experimental		/evidence-not_experimental
repeat_region	/rpt_family="MLT1A2"	repeat_region	/rpt_family="AT_rich"
	17596..17878		31384..31512
	/evidence-not_experimental		/evidence-not_experimental
repeat_region	/rpt_family="MLT2E"	repeat_region	/rpt_family="L2"
	complement(17884..18823)		32261..32348
	/evidence-not_experimental		/evidence-not_experimental
repeat_region	/rpt_family="L1"	repeat_region	/rpt_family="MIR"
	complement(19188..19990)		complement(33414..33633)
	/evidence-not_experimental		/evidence-not_experimental
repeat_region	/rpt_family="HERV16"	repeat_region	complement(33782..34282)

```

/evidence-not_experimental
/rpt_family="LIPA2"
complement(34414..34574)
repeat_region /evidence-not_experimental
/rpt_family="AluY"
35473..35652
/evidence-not_experimental
/rpt_family="MIR"
35689..36037
repeat_region /evidence-not_experimental
/rpt_family="MLT1"
36911..37100
/evidence-not_experimental
/rpt_family="MIR"
complement(37413..37458)
repeat_region /evidence-not_experimental
/rpt_family="L2"
complement(37459..37482)
repeat_region /evidence-not_experimental
/rpt_family="MIR"
complement(37650..37951)
repeat_region /evidence-not_experimental
/rpt_family="LIR50"
complement(37919..38131)

Alignment Scores:
Pred. No.: 542 Length: 183125
Score: 61.00 Matches: 12
Percent Similarity: 93.33% Conservative: 2
Best Local Similarity: 80.00% Mismatches: 1
Query Match: 81.33% Indels: 0
DB: 9 Gaps: 0

US-10-075-846-4_COPY_192_207 (1-16) x AP002848 (1-183125)
QY 1 SerSerSerIleuCySerProleuProSerleuSerleuSer 15
Db 149928 ACACATCAATTCGTGCTCACCTCCCAACCTCAGCTCTCT 149884

RESULT 5
AC084114/C
LOCUS AC084114 189127 bp DNA linear PRI 01-FEB-2002
DEFINITION Homo sapiens chromosome 8, clone RP11-1059L18, complete sequence.
ACCESSION AC084114
VERSION AC084114.8 GI:18464106
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 189127)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Beda,F., Boguslavsky,L.,
Bouhagalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
Chapel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,
Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M.,
Graham,I., Johnson,R., Jones,C., Kann,L., Heaford,A., Horton,L.,
Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., Larocque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,C., Liu,G.,
Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McKernan,K.,
McPheeters,R., Meldrim,J., Menus,L., Mihova,T., Mlenga,V.,
Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,
O'Donnell,P., O'Neill,D., Olivat,T.M., Oliver,J., Peterson,K.,
Pierre,N., Pisanu,C., Pollara,V., Raymond,C., Rieback,M., Riley,R.,
Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P.,
Souarez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,

```

```

TITLE Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
DIRECT Submision
JOURNAL Submitted (13-OCT-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 3 (bases 1 to 189127)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., All,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Bouhagalter,B.,
Brown,A., Camarato,J., Campopiano,A., Chang,J., Chazaro,B.,
Chapel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,A.,
Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,I., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamit,A., Karatas,A., Kells,C., Larocque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Liu,G., Maclean,C.,
Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
McEwan,P., McKernan,K., Meldrim,J., Menus,L., Mihova,T.,
Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
Rosetti,M., Roy,A., Santos,R., Schauer,S., Schnupack,R., Seaman,S.,
Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (01-FEB-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Feb 1, 2002 this sequence version replaced g1:18450000.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green,P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIRB
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L10840
Center clone name: 1059_L_18
-----
FEATURES
source
1..189127
location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8"
/clone="RP11-1059L18"
/clone_lib="RP11-11 Human Male BAC"
complement(14..319)
/rpt_family="MLT1"
113..117
/note="<30 qual SINGL region."
167..171
/note="<30 qual SINGL region."
1550..2067
/rpt_family="MLT1H1"
2213..2400
/rpt_family="MIR3"
complement(3205..3249)
/rpt_family="L2"
3250..3563
/rpt_family="AluX"
complement(3564..3608)
/rpt_family="L2"
complement(3927..4359)
/rpt_family="LIME1"
4481..4806
/rpt_family="L2"
4844..4974
repeat_region

```

```

repeatL_region /rpt_family="LIMB8"
5477. .5791
/rpt_family="ALUSX"
repeatL_region 6550. .6590
/rpt_family="(TG)n"
complement(6605. .6831)
/rpt_family="MLT1E2"
complement(6832. .7762)
/rpt_family="LIMC4"
complement(7763. .7980)
/rpt_family="LIM1"
complement(7981. .8070)
/rpt_family="LIMC4"
complement(8101. .8163)
/rpt_family="LIMC2"
complement(8159. .8340)
/rpt_family="LIMC3"
complement(8341. .8640)
/rpt_family="ALUSg"
complement(8641. .8967)
/rpt_family="LIMC3"
8968. .8993
/rpt_family="(TTTA)n"
complement(8994. .9235)
/rpt_family="LIMC3"
9233. .9440
/rpt_family="LIMC4a"
complement(9454. .9695)
/rpt_family="LIMC"
complement(9735. .9993)
/rpt_family="LIMC"
10007. .10042
/rpt_family="AT_rich"
complement(10411. .10663)
/rpt_family="MLT1K"
complement(10703. .11218)
/rpt_family="L2"
complement(11251. .11499)
/rpt_family="MLT1K"
complement(12557. .12621)
/rpt_family="L2"
12819. .13178
/rpt_family="L2"
13272. .13379
/rpt_family="L2"
complement(13423. .13599)
/rpt_family="L2"
complement(14738. .15236)
/rpt_family="MLT1U"
complement(15831. .15971)
/rpt_family="LTR27"
complement(16004. .16578)
/rpt_family="LTR27"
16590. .17021
/rpt_family="MLT1C"
17022. .17479
/rpt_family="L2"
17745. .17871
/rpt_family="L2"
18053. .18211
/rpt_family="MIR"
19166. .19187
/rpt_family="(T)n"
complement(19989. .20181)
/rpt_family="MIR"
20234. .20579
/rpt_family="LIM4"
21771. .21911
/rpt_family="MIR3"
22077. .22098
/rpt_family="AT_rich"
complement(22383. .22338)
/rpt_family="MLT1"

```

```

repeatL_region complement(22792. .22944)
/rpt_family="MIR"
23103. .23290
/rpt_family="MIR"
23294. .23337
/rpt_family="(CAT)n"
23737. .24025
/rpt_family="L2"
24164. .24244
/rpt_family="LTR33A"
24434. .24517
/rpt_family="L2"
complement(24797. .25032)
/rpt_family="L2"
complement(27098. .27304)
/rpt_family="MIR"
27711. .27780
/rpt_family="(TATATG)n"
complement(28712. .28808)
/rpt_family="L2"
complement(28810. .29053)
/rpt_family="LIM3"

```

```

Alignment Scores:
Pred. No.: 560 Length: 189127
Score: 61.00 Matches: 12
Percent Similarity: 93.33% Conservative: 2
Best Local Similarity: 80.008 Mismatches: 1
Query Match: 81.338 Indels: 0
DB: 9 Gaps: 0

```

US-10-075-846-4\_COPY\_192\_207 (1-16) x AC084114 (1-189127)

```

QY 1 SerSerSerIleucySerProIeuProSerIeuSerIeuSer 15
::: |||||
Db 36618 ACACATCAATCTGCTGCTACCTCCACCACTCAGCTCTCT 36574

```

```

RESULT 6
LOCUS AC079196 195477 bp DNA linear HTG 18-NOV-2000
DEFINITION Homo sapiens chromosome 8 clone RP11-480F5 map 8, WORKING DRAFT
ACCESSION AC079196
VERSION AC079196.2 GI:11225390
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 195477)
Homo sapiens chromosome 8, clone RP11-480F5
Unpublished
2 (bases 1 to 195477)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Anderson,S., Barna,N., Bastien,V., Beda,F., Boguslavskiy,L.,
Boukhalter,B., Brown,A., Burkett,C., Campopiano,A., Castle,A.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,
Fitzhugh,W., Gage,D., Galagan,J., Gargyala,S., Ginde,S., Coyette,M.,
Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L.,
Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., LaRocque,K.,
Lamazares,R., Landers,T., Lahoczky,J., Levine,R., Liu,C., Liu,G.,
Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McKernan,K.,
McPheters,R., Meldrim,J., Meneus,L., Mihova,T., Mieng,V.,
Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,
O'Donnell,P., O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K.,
Pierre,N., Pisanic,C., Pollara,V., Raymond,C., Ribbeck,M., Riley,R.,
Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P.,
Sougnuez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Stauss,N., Subramanian,A., Talamas,J., Testaiey,S., Theodore,J.,
Tirelli,A., Travers,M., Trifilio,J., Vassiliev,H., Viel,R., Vo,A.,
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,

```

TITLE  
JOURNAL  
COMMENT

Zimmer, A. and Zody, M.  
Direct Submission  
Submitted (23-AUG-2000) Whitehead Institute/MIT Center for Genome  
Research, 330 Charles Street, Cambridge, MA 02141, USA  
On Nov 18, 2000 this sequence version replaced g1:9886032.  
All repeats were identified using RepeatMasker:  
Shtlt, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WtBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence\_submissions@genome.wi.mit.edu

Project Information

Center project name: L10608

Center clone name: 480.F.5

Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 180424 bases at least Q40  
Consensus quality: 188415 bases at least Q30  
Consensus quality: 191442 bases at least Q20  
Insert size: 198000; agarose-ff  
Insert size: 193277; sum-of-ctrls  
Quality coverage: 4.2 in Q20 bases; agarose-ff  
Quality coverage: 4.3 in Q20 bases; sum-of-ctrls

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 23 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1  
776 875: contig of 775 bp in length  
776 875: gap of 100 bp  
876 1876: contig of 1001 bp in length  
1877 1976: gap of 100 bp  
1977 58986: contig of 57010 bp in length  
58987 59086: gap of 100 bp  
59087 60286: contig of 1203 bp in length  
60286 60389: gap of 100 bp  
60390 61513: contig of 1124 bp in length  
61514 61613: gap of 100 bp  
61614 62903: contig of 1290 bp in length  
62904 63003: gap of 100 bp  
63004 64788: contig of 1785 bp in length  
64789 64888: gap of 100 bp  
64889 67987: contig of 3099 bp in length  
67988 68087: gap of 100 bp  
68088 71902: contig of 3815 bp in length  
71903 72002: gap of 100 bp  
72003 74471: contig of 2465 bp in length  
74472 74571: gap of 100 bp  
74572 77764: contig of 3193 bp in length  
77765 77864: gap of 100 bp  
77865 81116: contig of 3252 bp in length  
81117 81216: gap of 100 bp  
81217 85155: contig of 3939 bp in length  
85156 85255: gap of 100 bp  
85256 89134: contig of 3879 bp in length  
89135 89234: gap of 100 bp  
89235 96530: contig of 7296 bp in length  
96531 96630: gap of 100 bp  
96631 104817: contig of 8187 bp in length  
104818 104917: gap of 100 bp  
104918 118601: contig of 13684 bp in length  
118602 118701: gap of 100 bp  
118702 129776: contig of 11075 bp in length  
129777 129876: gap of 100 bp  
129877 140248: contig of 10372 bp in length

# FEATURES

140249 140348: gap of 100 bp  
140349 150907: contig of 10559 bp in length  
150908 151007: gap of 100 bp  
151008 168203: contig of 18196 bp in length  
168204 169303: gap of 100 bp  
169304 194921: contig of 25618 bp in length  
194922 195021: gap of 100 bp  
195022 195477: contig of 456 bp in length.  
Location/Qualifiers

1. 195477  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="8"  
/map="8"  
/clone="RP11-480P5"  
/clone\_11b="RPCT-11 Human Male BAC"  
1. .775  
/note="assembly-fragment  
clone\_end:SP6  
vector\_side:left"  
876. .1876  
/note="assembly-fragment"  
1977. .58986  
/note="assembly-fragment"  
59087. .60289  
/note="assembly-fragment"  
60390. .61513  
/note="assembly-fragment"  
61614. .62903  
/note="assembly-fragment"  
63004. .64788  
/note="assembly-fragment"  
64889. .67987  
/note="assembly-fragment"  
68088. .71902  
/note="assembly-fragment"  
72003. .74471  
/note="assembly-fragment"  
74572. .77764  
/note="assembly-fragment"  
77865. .81116  
/note="assembly-fragment"  
81217. .85155  
/note="assembly-fragment"  
85256. .89134  
/note="assembly-fragment"  
89235. .96530  
/note="assembly-fragment"  
96631. .104817  
/note="assembly-fragment"  
104918. .118601  
/note="assembly-fragment"  
118702. .129776  
/note="assembly-fragment"  
129877. .140248  
/note="assembly-fragment"  
140349. .150907  
/note="assembly-fragment"  
151008. .169203  
/note="assembly-fragment"  
169304. .194921  
/note="assembly-fragment"  
195022. .195477  
/note="assembly-fragment  
clone\_end:T7  
vector\_side:right"

BASE COUNT 58603 a 36972 c 36289 g 61402 t 2211 others  
ORIGIN

Alignment Scores: 579 Length: 195477  
Pred. No.: 61.00 Matches: 12  
Score: 93.33% Conservative: 2  
Percent Similarity:

Best Local Similarity: 80.00%  
Query Match: 81.33%  
DB: 2 Gaps: 0  
US-10-075-846-4\_COPY\_192\_207 (1-16) x AC079196 (1-195477)  
Oy 1 SerSerSerIleLeuGysSerProLeuProSerLeuSerLeuSer 15  
Db 46364 ACACATCATTCATTCCTCCTCCTCCCAACCTCAGCTCTCT 46408  
RESULT 7  
AC109887  
LOCUS  
DEFINITION  
AC109887 91582 bp DNA linear HTG 13-JUL-2002  
Rattus norvegicus clone CH230-312G3, \*\*\* SEQUENCING IN PROGRESS  
\*\*\*, 49 unordered pieces.  
AC109887 3 GI:21738183  
AC109887.3  
HTG: HTGS.PHASE1.  
Norway rat.  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 91582)  
Munzy,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,  
Alsbrooks,S.L., Amarantunga,H.C., Are,J.R., Ayele,M., Banks,T.,  
Barbata,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,  
Bouck,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,  
Bunay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C.,  
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
David,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
DeJangey,K.R., Delgado,O., Deon,A.L., Ding,X., Dinh,H.H.,  
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
Eamhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
Gorell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,  
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,  
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,  
Honsi,F., Howard,S., Huber,J., Huliy,S., Hume,J., Jackson,L.E.,  
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
Karlsom,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovai,C.,  
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
Li,J., Li,Z., Licharge,O., Lieu,C., Liu,J., Liu,W., Louised,H.,  
Lorado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,  
Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mei,G., Metker,M.,  
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,  
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,  
Nguyen,N., Nickerson,E., Nwokoko,S., Ogub,M., Okunolu,G.,  
Orangunye,N., Oyiedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
Peterson,R., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,  
Rives,M., Rojas,A., Rojebokan,I., Rolfe,M., Ruiz,S., Savary,G.,  
Scherer,S., Scott,G., Shen,H., Shoshitaishvili,N., Stinson,I.,  
Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,  
Sutton,A., Swalek,A., Tabor,P., Tameris,A., Tameris,K., Tang,H.,  
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
Usmani,K., Vasquez,U., Vera,V., Villalón,D., Vinson,R., Wang,Q.,  
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,  
Williams,G., Williamson,A., Wleczek,R., Woodson,S., Worley,K.,  
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorilla,S., Nelson,D.,  
Weinstock,G. and Gibbs,R.  
Direct Submission  
Unpublished  
2 (bases 1 to 91582)  
Worley,K.C.  
JOURNAL  
AUTHORS  
REFERENCE  
Submitted (08-FEB-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 91582)

AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Worley,K.C.  
Direct Submission  
Submitted (13-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Jul 12, 2002 this sequence version replaced gi:18860208.  
-----  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: http://www.hgsc.bcm.tmc.edu/  
Contact: hgsc-help@bcm.tmc.edu  
-----  
Project Information  
Center project name: GOKI  
Center clone name: CH230-312G3  
-----  
Summary Statistics  
Sequencing vector: plasmid  
Chemistry: Dye-terminator Big Dye 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 62571 bases at least Q40  
Consensus quality: 66037 bases at least Q30  
Consensus quality: 69135 bases at least Q20  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 49 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
1  
1503 1502: contig of 1502 bp in length  
1803 1502: gap of unknown length  
2721 1502: contig of 1119 bp in length  
2821 1502: gap of unknown length  
3963 1502: contig of 1142 bp in length  
3964 1502: gap of unknown length  
4063 1502: contig of 1042 bp in length  
5106 1502: gap of unknown length  
5206 1502: contig of 1372 bp in length  
6578 1502: gap of unknown length  
6677 1502: gap of 1125 bp in length  
7802 1502: contig of 1044 bp in length  
7903 1502: gap of unknown length  
9047 1502: gap of 1071 bp in length  
10117 1502: contig of 1071 bp in length  
10217 1502: gap of unknown length  
10218 1502: contig of 1262 bp in length  
11479 1502: gap of unknown length  
11480 1502: gap of 1265 bp in length  
11844 1502: contig of 1265 bp in length  
12845 1502: gap of unknown length  
12945 1502: contig of 2006 bp in length  
14951 1502: gap of unknown length  
15051 1502: gap of 1592 bp in length  
16642 1502: contig of 1592 bp in length  
16742 1502: gap of unknown length  
17828 1502: contig of 1086 bp in length  
17829 1502: gap of unknown length  
17928 1502: contig of 1799 bp in length  
19727 1502: gap of unknown length  
19827 1502: gap of 1536 bp in length  
21363 1502: contig of 1536 bp in length  
21463 1502: gap of unknown length  
22652 1502: contig of 1189 bp in length  
22752 1502: gap of unknown length  
22753 1502: contig of 1177 bp in length  
23930 1502: gap of unknown length  
24029 1502: gap of 1549 bp in length  
25578 1502: contig of 1549 bp in length  
25678 1502: gap of unknown length  
26962 1502: contig of 1284 bp in length  
27062 1502: gap of unknown length  
27063 1502: contig of 1410 bp in length  
28472 1502: gap of unknown length  
28572 1502: gap of unknown length  
30264 1502: contig of 1692 bp in length



30265 30364: gap of unknown length  
30365 31464: contig of 1100 bp in length  
31465 31564: gap of unknown length  
31565 33215: contig of 1651 bp in length  
33216 33315: gap of unknown length  
33316 34890: contig of 1575 bp in length  
34891 34990: gap of unknown length  
34991 36624: contig of 1634 bp in length  
36625 36724: gap of unknown length  
36725 38529: contig of 1805 bp in length  
38530 38630: gap of unknown length  
38630 40251: contig of 1622 bp in length  
40252 40351: gap of unknown length  
40352 42287: contig of 1936 bp in length  
42288 42387: gap of unknown length  
42388 43463: contig of 1076 bp in length  
43464 43563: gap of unknown length  
43564 44922: contig of 1355 bp in length  
44923 45022: gap of unknown length  
45023 47349: contig of 2327 bp in length  
47350 47449: gap of unknown length  
47450 48918: contig of 1469 bp in length  
48919 49018: gap of unknown length  
49019 51017: contig of 1999 bp in length  
51018 51117: gap of unknown length  
51118 52930: contig of 1813 bp in length  
52931 53031: gap of unknown length  
53031 54539: contig of 1509 bp in length  
54540 54639: gap of unknown length  
54640 57214: contig of 2575 bp in length  
57215 57314: gap of unknown length  
57315 59204: contig of 1890 bp in length  
59205 59304: gap of unknown length  
59305 60829: contig of 1525 bp in length  
60830 60929: gap of unknown length  
60930 63121: contig of 2192 bp in length  
63122 63221: gap of unknown length  
63222 65614: contig of 2393 bp in length  
65615 65714: gap of unknown length  
65715 67972: contig of 2258 bp in length  
67973 68072: gap of unknown length  
68073 69880: contig of 1808 bp in length  
69881 69980: gap of unknown length  
69981 73676: contig of 3656 bp in length  
73677 73776: gap of unknown length  
73777 76126: contig of 2350 bp in length  
76127 76226: gap of unknown length  
76227 78822: contig of 2596 bp in length  
78823 78922: gap of unknown length  
78923 82017: contig of 3095 bp in length  
82018 82117: gap of unknown length  
82118 84115: contig of 1998 bp in length  
84116 84215: gap of unknown length  
84216 87398: contig of 3183 bp in length  
87399 87498: gap of unknown length  
87499 91582: contig of 4084 bp in length.  
Location/Qualifiers  
1. 91582  
/organism="Rattus norvegicus"  
/db\_xref="taxon:10116"  
/clone="CH230-31263"

BASE COUNT 24682 a 18454 c 18294 g 25285 t 4867 others  
ORIGIN

Alignment Scores:  
Pred. No.: 387  
Score: 60.00  
Percent Similarity: 92.868  
Best Local Similarity: 85.714  
Query Match: 80.004  
DB: 2  
Length: 91582  
Matches: 12  
Conservative: 1  
Mismatch: 1  
Indels: 0  
Gaps: 0

US-10-075-846-4\_COPY\_192\_207 (1-16) x AC109887 (1-91582)

OY 2 SerSer1leucysSerProleuProserLeuSerleuSer 15  
Db 65489 TCCAGCATCTCTCTCTCTCCATTTCCTACACGTCTCC 65530  
RESULT 8  
AC095732/c  
LOCUS  
DEFINITION  
Rattus norvegicus clone CH230-9A18, \*\*\* SEQUENCING IN PROGRESS \*\*\*,  
67 unoriented pieces.

AC095732 169845 bp DNA linear HTG 11-JUL-2002  
AC095732  
AC095732.4 GI:21722742  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
HTG: HTGS.PHASE1.  
Norway rat.  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE  
1 (bases 1 to 169845)  
Muzny,D.M., Adams,C., Adio-Oduola,B., All-osman,F.R., Allen,C.,  
Alsbrooks,S.L., Amarantunge,H.C., Are,J.R., Ayala,M., Banks,T.,  
Barberia,J., Benton,J., Bimge,K., Blankenburg,K., Bonnin,D.,  
Bouck,J., Bowls,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,  
Bunay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
Carton,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,  
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
Devalla,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,  
Doutwaite,K.J., Draper,H., Dugan-Rocha,S., Durin,K.J.,  
Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
Gablis,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,  
Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J.,  
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,  
Homst,F., Howard,S., Huber,J., Huiyk,S., Hume,J., Jackson,L.E.,  
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
Karlssoen,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,  
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
Li,J., Li,Z., Lichtarge,O., Lileu,C., Liu,J., Liu,W., Louisedge,H.,  
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,  
Massey,E., Mahoney,E., McLeod,M.P., Meador,M., Mel,G., Metzger,M.,  
Miner,G., Miner,Z., Mitchell,T., Mohabeh,K., Morgan,M., Morris,S.,  
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,  
Nguyen,N., Nickerson,E., Nockenwo,S., Oguh,M., Okuwonu,G.,  
Oragunye,N., Oyiedu,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,  
Rivers,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Saverly,C.,  
Scherer,S., Scott,G., Shen,H., Shooshari,N., Sisson,I.,  
Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,  
Sutton,A., Swatek,A., Tabori,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
Tansley,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
Umsani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q.,  
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,  
Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,  
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
Weinstock,G. and Gibbs,R.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
DIRECT Submission  
Unpublished  
2 (bases 1 to 169845)  
Worley,K.C.  
Direct Submission  
Submitted (17-SEP-2001) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 169845)  
Worley,K.C.  
Direct Submission  
Submitted (11-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

## COMMENT

On Jul 10, 2002 this sequence version replaced gi:20975871.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GDEZ

Center clone name: CH230-9A18

----- Summary Statistics

Sequencing vector: Plasmid

Chemistry: Dye-terminator Big Dye: 100% of reads

Assembly program: Phrap: version 0.990329

Consensus quality: 104990 bases at least Q40

Consensus quality: 112511 bases at least Q30

Consensus quality: 119479 bases at least Q20

----- NOTE: Estimated insert size may differ from sequence length

NOTE: (see [http://www.hgsc.bcm.tmc.edu/docs/genbank/draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html)).

NOTE: This is a 'working draft' sequence. It currently

consists of 67 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence.

as soon as it is available and the accession number will

be preserved.

1 1038: contig of 1038 bp in length

1039 1138: gap of unknown length

1139 2221: contig of 1083 bp in length

2222 2321: gap of unknown length

2322 3547: contig of 1226 bp in length

3548 3647: gap of unknown length

3648 4983: contig of 1336 bp in length

4984 5083: gap of unknown length

5084 6222: contig of 1139 bp in length

6223 7585: gap of unknown length

7586 7685: contig of 1263 bp in length

7686 8858: gap of unknown length

8859 10012: contig of 1173 bp in length

10013 10112: gap of unknown length

10113 11455: contig of 1343 bp in length

11456 11555: gap of unknown length

11556 13017: contig of 1462 bp in length

13018 13117: gap of unknown length

13118 14720: contig of 1603 bp in length

14721 14820: gap of unknown length

14821 16158: contig of 1338 bp in length

16159 16258: gap of unknown length

16259 17750: contig of 1492 bp in length

17751 17850: gap of unknown length

17851 19465: contig of 1615 bp in length

19466 19565: gap of unknown length

19566 20786: contig of 1221 bp in length

20787 20886: gap of unknown length

20887 22283: contig of 1397 bp in length

22284 22383: gap of unknown length

22384 23550: contig of 1167 bp in length

23551 25247: contig of 1597 bp in length

25248 25347: gap of unknown length

25348 26739: contig of 1392 bp in length

26740 26839: gap of unknown length

26840 28297: contig of 1458 bp in length

28298 28397: gap of unknown length

28398 30539: contig of 2142 bp in length

30540 30639: gap of unknown length

30640 31984: contig of 1345 bp in length

31985 32084: gap of unknown length

32085 33203: contig of 1119 bp in length

33204 33303: gap of unknown length

33304 35105: contig of 1802 bp in length

35106 35205: gap of unknown length

35206 35554: contig of 1349 bp in length

35555 36554: gap of unknown length

36556 39009: contig of 2355 bp in length

39010 39109: gap of unknown length

39110 40559: contig of 1450 bp in length

40560 40659: gap of unknown length

40660 41768: contig of 1109 bp in length

41769 41868: gap of unknown length

41869 44085: contig of 2217 bp in length

44086 44185: gap of unknown length

44186 46351: contig of 2166 bp in length

46352 46451: gap of unknown length

46452 48980: contig of 2529 bp in length

48981 49080: gap of unknown length

49081 51093: contig of 2013 bp in length

51094 52710: contig of 1517 bp in length

52711 52810: gap of unknown length

52811 55632: contig of 2822 bp in length

55633 55732: gap of unknown length

55733 58525: contig of 2793 bp in length

58526 61423: gap of unknown length

61424 61523: contig of 2798 bp in length

61524 63441: gap of unknown length

63442 63541: gap of unknown length

63542 66215: contig of 2674 bp in length

66216 66315: gap of unknown length

66316 68159: contig of 1844 bp in length

68160 70262: gap of unknown length

70263 70362: contig of 2003 bp in length

70363 71758: gap of unknown length

71759 71858: contig of 1396 bp in length

71859 73318: gap of unknown length

73319 73418: contig of 1460 bp in length

73419 75122: gap of unknown length

75123 75222: contig of 1704 bp in length

75223 78668: gap of unknown length

78669 81040: contig of 3446 bp in length

81041 81140: gap of unknown length

81141 83306: contig of 2272 bp in length

83307 83406: gap of unknown length

83407 86465: contig of 2166 bp in length

86466 86565: gap of unknown length

86566 89067: contig of 3059 bp in length

89068 89167: gap of unknown length

89168 92835: contig of 2502 bp in length

92836 92935: gap of unknown length

92936 95506: contig of 3668 bp in length

95507 95606: gap of unknown length

95607 98304: contig of 2571 bp in length

98305 98404: gap of unknown length

98405 101381: contig of 2698 bp in length

101382 101481: gap of unknown length

101481 105478: contig of 2976 bp in length

105479 105578: gap of unknown length

105579 108776: contig of 3198 bp in length

## Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-10-075-846-4\_COPY\_192\_207 (1-16) x AC095732 (1-169845)

Oy

1 SerSerSerIleLeucySerProLeuProSerLeuSerLeuSer 15

2.07e+03 Length: 169845

57.00 Matches: 13

86.67% Conservative: 0

86.67% Mismatches: 2

76.00% Indels: 0

2 Gaps: 0





Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GDOJ  
Center clone name: CH230-9122  
----- Summary Statistics  
Sequencing vector: Plasmid  
Chemistry: Dye-terminator Big Dye 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 44880 bases at least Q40  
Consensus quality: 54834 bases at least Q30  
Consensus quality: 62109 bases at least Q20  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank/draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 60 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
1 1024: contig of 1024 bp in length  
\* 1 1025 1124: gap of unknown length  
\* 1125 2135: contig of 1011 bp in length  
\* 2135 2235: gap of unknown length  
\* 2235 3600: contig of 1365 bp in length  
\* 3600 3700: gap of unknown length  
\* 3700 4798 4897: gap of unknown length  
\* 4798 4898 5991: contig of 1094 bp in length  
\* 4898 5992 6092: gap of unknown length  
\* 6092 7247 7346: contig of 1155 bp in length  
\* 7247 7347 8487: gap of unknown length  
\* 7347 8488 8587: gap of unknown length  
\* 8488 8588 9639: contig of 1052 bp in length  
\* 8588 9640 9739: gap of unknown length  
\* 9640 11091 11090: contig of 1351 bp in length  
\* 11091 11191 11190: gap of unknown length  
\* 11191 12196: contig of 1006 bp in length  
\* 12196 12997 13886: gap of unknown length  
\* 12997 13887 13986: gap of unknown length  
\* 13887 13987 15685: contig of 1699 bp in length  
\* 13987 15685 15785: gap of unknown length  
\* 15685 15785 17641: contig of 1856 bp in length  
\* 15785 17641 17741: gap of unknown length  
\* 17641 20288 20388: contig of 2347 bp in length  
\* 20288 20388 21497: gap of unknown length  
\* 20388 21497 21597: contig of 1109 bp in length  
\* 21497 21597 22668: gap of unknown length  
\* 21597 22668 22768: contig of 1071 bp in length  
\* 22668 22768 24455: gap of unknown length  
\* 22768 24455 24555: contig of 1687 bp in length  
\* 24455 24555 26378: gap of unknown length  
\* 24555 26378 27628: contig of 1250 bp in length  
\* 26378 27628 29350: gap of unknown length  
\* 27628 29350 29350: contig of 1522 bp in length  
\* 29350 29350 30860: gap of unknown length  
\* 29350 30860 32051: contig of 1510 bp in length  
\* 30860 32051 32151: gap of unknown length  
\* 32051 32151 34551: contig of 1091 bp in length  
\* 32151 34551 34651: gap of unknown length  
\* 34551 34651 35991: contig of 1340 bp in length  
\* 34651 35991 36091: gap of unknown length  
\* 35991 36091 37109: contig of 1018 bp in length  
\* 36091 37109 37209: gap of unknown length  
\* 37109 37209 37209: gap of unknown length

37210 38947: contig of 1738 bp in length  
\* 38947 39047: gap of unknown length  
\* 39047 40168: contig of 1121 bp in length  
\* 40168 40268: gap of unknown length  
\* 40268 41542: contig of 1274 bp in length  
\* 41542 41642: gap of unknown length  
\* 41642 43437: contig of 1795 bp in length  
\* 43437 43537: gap of unknown length  
\* 43537 45119: contig of 1582 bp in length  
\* 45119 45120: gap of unknown length  
\* 45120 46865: contig of 1646 bp in length  
\* 46865 46965: gap of unknown length  
\* 46965 48123: contig of 1158 bp in length  
\* 48123 48224: gap of unknown length  
\* 48224 49481: contig of 1288 bp in length  
\* 49481 49581: gap of unknown length  
\* 49581 50995: contig of 1414 bp in length  
\* 50995 51095: gap of unknown length  
\* 51095 52439: contig of 1344 bp in length  
\* 52439 52539: gap of unknown length  
\* 52539 54379: contig of 1740 bp in length  
\* 54379 54380: gap of unknown length  
\* 54380 56335: contig of 2256 bp in length  
\* 56335 56735: gap of unknown length  
\* 56735 58377: contig of 1642 bp in length  
\* 58377 58477: gap of unknown length  
\* 58477 59989: contig of 1512 bp in length  
\* 59989 60089: gap of unknown length  
\* 60089 61584: contig of 1495 bp in length  
\* 61584 61684: gap of unknown length  
\* 61684 62912: contig of 1228 bp in length  
\* 62912 63012: gap of unknown length  
\* 63012 64579: contig of 1567 bp in length  
\* 64579 64580: gap of unknown length  
\* 64580 66134: contig of 1455 bp in length  
\* 66134 66234: gap of unknown length  
\* 66234 68459: contig of 2225 bp in length  
\* 68459 68559: gap of unknown length  
\* 68559 69792: contig of 1233 bp in length  
\* 69792 69892: gap of unknown length  
\* 69892 71547: contig of 1655 bp in length  
\* 71547 71548 71647: gap of unknown length  
\* 71647 73376: contig of 1729 bp in length  
\* 73376 73476: gap of unknown length  
\* 73476 76099: contig of 2623 bp in length  
\* 76099 76100 78499: gap of unknown length  
\* 78499 78646: contig of 2447 bp in length  
\* 78646 78746 80040: gap of unknown length  
\* 80040 80140: gap of unknown length  
\* 80140 82556: contig of 2416 bp in length  
\* 82556 82656: gap of unknown length  
\* 82656 84373: contig of 1617 bp in length  
\* 84373 84374 86987: gap of unknown length  
\* 84374 86987 87087: contig of 2614 bp in length  
\* 87087 87088 89267: gap of unknown length  
\* 89267 89267: contig of 2180 bp in length

## Alignment Scores:

Pred. No.: 1 82e+03 Length: 104525  
Score: 56.00 Matches: 11  
Percent Similarity: 86.67% Conservative: 2  
Best Local Similarity: 73.33% Mismatches: 2  
Query Match: 74.67% Indels: 0  
DB: 2 Gaps: 0

US-10-075-846-4\_COPY\_192\_207 (1-16) x AC095220 (1-104525)

Qy 1 SerSerSeriLeuCySerProLeuProSerLeuSerLeuSer 15  
||||| :|||||:|||||:|||||:|||||:|||||:|||||  
Db 43411 TCATCGGTTCTCTCTGCTCCCTCTCTTCAATCACTGCTCC 43367

RESULT 11

AC102866 172555 bp DNA linear HTG 20-AUG-2002  
 LOCUS Mus musculus clone RP24-576L9, WORKING DRAFT SEQUENCE, 12 unordered  
 DEFINITION pieces.  
 AC102866  
 AC102866.2 GI:22325064  
 VERSION HTG; HTGS\_PHASE1; HTGS\_DRAFT; HMGS-FULLTOP.  
 KEITHOROS house mouse.  
 SOURCE Mus musculus.  
 ORGANISM Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 172555)  
 AUTHORS Birren,B., Nusbaum,C. and Lander,E.  
 TITLE Mus musculus, clone RP24-576L9  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 172555)  
 AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., All,A., Allen,N.,  
 Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhalter,B.,  
 Brown,A., Camarata,J., Campolano,A., Chang,J., Chazaro,B.,  
 Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,  
 Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,  
 Ferreira,P., Fitzhugh,M., Gage,D., Galagan,J., Gardyna,S.,  
 Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,  
 Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,  
 Jones,C., Kamat,A., Karatas,A., Kells,C., LaRoque,R.,  
 Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,  
 McCarty,M., McEwan,P., McKernan,K., MCPheeters,R., Meldrum,J.,  
 Menus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,  
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,  
 Oliver,J., Peterson,K., Pihunkhang,P., Pierre,N., Pollard,V.,  
 Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,  
 Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,N., Schupack,R.,  
 Seaman,S., Severy,P., Spencer,B., Strange-Thomas,N., Stojanovic,N.,  
 Strauss,K., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,  
 Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,  
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.-J., Young,G.,  
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.  
 JOURNAL Direct Submission  
 REFERENCE Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome  
 Research, 330 Charles Street, Cambridge, MA 02141, USA  
 AUTHORS 3 (bases 1 to 172555)  
 REFERENCE Birren,B., Nusbaum,C., Lander,E., All,A., Allen,N., Anderson,S.,  
 Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B.,  
 Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,  
 Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,  
 Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,  
 Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,  
 Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,  
 Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,  
 Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,  
 McCarty,M., Meldrum,J., Menus,L., Mihova,T., Mlenga,V.,  
 Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,  
 O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,  
 Pihunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,  
 Roman,J., Roy,A., Schauer,S., Schupack,R., Seaman,S., Severy,P.,  
 Smith,C., Spencer,B., Strange-Thomas,N., Stojanovic,N., Talamas,J.,  
 Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,  
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,  
 Zembek,L., Zimmer,A. and Zody,M.  
 JOURNAL Direct Submission  
 COMMENT Submitted (20-AUG-2002) Whitehead Institute/MIT Center for Genome  
 Research, 330 Charles Street, Cambridge, MA 02141, USA  
 On Aug 20, 2002 this sequence version replaced gi:17060976.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html  
 ----- Genome Center  
 Center: Whitehead Institute/MIT Center for Genome Research  
 Center code: MIBR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 ----- Project Information

Center project name: L19975  
 Center Clone name: 576\_L-9  
 ----- Summary Statistics  
 Sequencing vector: Plasmid; n/a; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.960731  
 Consensus quality: 169439 bases at least Q40  
 Consensus quality: 170530 bases at least Q30  
 Consensus quality: 170959 bases at least Q20  
 Insert size: 157000; agarose-fp  
 Insert size: 171455; sum-of-contigs  
 Quality coverage: 12.8 in Q20 bases; agarose-fp  
 Quality coverage: 11.7 in Q20 bases; sum-of-contigs  
 -----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 12 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \*  
 \* 1  
 \* 635 734: gap of 100 bp  
 \* 735 1427: contig of 693 bp in length  
 \* 1428 1527: gap of 100 bp  
 \* 1528 2188: contig of 661 bp in length  
 \* 2189 2288: gap of 100 bp  
 \* 2289 2980: contig of 692 bp in length  
 \* 2981 3080: gap of 100 bp  
 \* 3081 3718: contig of 638 bp in length  
 \* 3719 3818: gap of 100 bp  
 \* 3819 4462: contig of 644 bp in length  
 \* 4463 4562: gap of 100 bp  
 \* 4563 5222: contig of 660 bp in length  
 \* 5223 5322: gap of 100 bp  
 \* 5323 6387: contig of 1065 bp in length  
 \* 6388 6487: gap of 100 bp  
 \* 6488 8217: contig of 1730 bp in length  
 \* 8218 8317: gap of 100 bp  
 \* 8318 9778: contig of 1461 bp in length  
 \* 9779 9878: gap of 100 bp  
 \* 9879 89723: contig of 79845 bp in length  
 \* 89724 89824: gap of 100 bp  
 \* 89824 172555: contig of 82732 bp in length.  
 \*  
 \* Location/Qualifiers  
 \* 1. 172555  
 \* /organism="Mus musculus"  
 \* /db\_xref="taxon:10090"  
 \* /clone="RP24-576L9"  
 \* /clone\_lib="RPCT-24 Male Mouse BAC"  
 \* 1. 634  
 \* /note="assembly-fragment"  
 \* 735. 1427  
 \* /note="assembly-fragment"  
 \* 1528. 2188  
 \* /note="assembly-fragment"  
 \* 2289. 2980  
 \* /note="assembly-fragment"  
 \* 3081. 3718  
 \* /note="assembly-fragment"  
 \* 3819. 4462  
 \* /note="assembly-fragment"  
 \* 4563. 5222  
 \* /note="assembly-fragment"  
 \* 5323. 6387  
 \* /note="assembly-fragment"  
 \* 6488. 8217  
 \* /note="assembly-fragment"  
 \* 8318. 9778  
 \* /note="assembly-fragment"  
 \* 9879. 89723  
 \* /note="assembly-fragment"

misc\_feature 89824..172555  
/note="assembly-fragment"  
BASE COUNT 53892 a 31196 c 31503 g 54857 t 1107 others  
ORIGIN

Alignment Scores:  
Pred. No.: 3e+03 Length: 172555  
Score: 56.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 3  
Best Local Similarity: 76.92% Mismatches: 0  
Query Match: 74.67% Indels: 0  
DB: 2 Gaps: 0

US-10-075-846-4\_COPY\_192\_207 (1-16) x AC102866 (1-172555)

QY 4 ILLeucySerProLeuProSerLeuSerLeuSerVal 16  
Db 52101 TTGCTTCTCTCTCCCTTCCTCCATGTCCTCTCTCTT 52139

RESULT 12  
GG126MRP 598 bp mRNA linear VRT 28-APR-1992  
LOCUS G.Gallus 126 MRP mRNA.  
X61200  
VERSION X61200.1 GI:62998  
KEYWORDS  
SOURCE Chicken.  
ORGANISM Gallus gallus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
Phasianinae; Gallus.  
REFERENCE 1 (bases 1 to 598)  
AUTHORS Graf, T.H.  
TITLE Direct Submission  
JOURNAL Submitted (01-AUG-1991) T.H. Graf, EMBL, Meyerhofstr. 1, 6900  
Heidelberg, FRG  
2 (bases 1 to 598)  
AUTHORS Nakano, T. and Graf, T.  
TITLE Identification of genes differentially expressed in two types of  
v-myb-transformed avian myelomonocytic cells  
JOURNAL Oncogene 7 (3), 527-534 (1992)  
MEDLINE 92195690  
PUBMED 1549365

FEATURES  
source Location/Qualifiers  
1..598  
/organism="Gallus gallus"  
/db\_xref="taxon:9031"  
/clone="126"  
/rname="lib" rna="AMV(4106) promyelocytes"  
BASE COUNT 158 a 198 c 129 g 113 t  
ORIGIN

Alignment Scores:  
Pred. No.: 15.1 Length: 598  
Score: 55.00 Matches: 11  
Percent Similarity: 81.25% Conservative: 2  
Best Local Similarity: 68.75% Mismatches: 3  
Query Match: 73.33% Indels: 0  
DB: 5 Gaps: 0

US-10-075-846-4\_COPY\_192\_207 (1-16) x GG126MRP (1-598)

QY 1 SerSerSerIleLeuCySerProLeuProSerLeuSerLeuSerVal 16  
Db 482 TCTCTCTCAACCCCTCTCTCCCTCCCAACCCCTCTCTCATC 529

RESULT 13  
AC118327 139754 bp DNA linear HTG 14-JUL-2002  
LOCUS Rattus norvegicus clone CH230-501B21, \*\*\* SEQUENCING IN PROGRESS  
DEFINITION  
AC118327  
AC118327.2 GI:21747379

KEYWORDS HTG; HTGS\_PHASE1.  
SOURCE Norway rat.  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE 1 (bases 1 to 139754)  
AUTHORS Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C.,  
Albrooks, S.L., Amaralunga, H.C., Are, J.R., Ayala, M., Banks, T.,  
Barbata, U., Benton, J., Bimaga, K., Blankenburg, K., Bonini, D.,  
Bouck, J., Bowle, S., Brileva, M., Brown, M., Brown, M., Bryant, N.P.,  
Bunyah, C., Burck, P., Burkett, C., Burrell, K.L., Byrd, N.C.,  
Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,  
Chen, G., Chen, R., Chen, Z., Chowdhury, I., Christopoulos, C.,  
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,  
Dayila, M.L., Davis, C., Davy-Carroll, L., Dedert, D.A.,  
Delaney, K.R., Delgado, O., Dem, A.L., Ding, Y., Dinh, H.H.,  
Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,  
Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotlo, M.,  
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,  
Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,  
Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,  
Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,  
Hernandez, O., Hodgson, A., Hogues, M., Hollway, C., Hollins, B.,  
Homsel, F., Howard, S., Huber, J., Hulik, S., Hume, J., Jackson, L.E.,  
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,  
Karlssohn, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,  
Kratoch, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,  
Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louised, H.,  
Lozado, R.J., Lu, X., Luchter, A., Lucier, R., Luna, R., Ma, J.,  
Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,  
Massey, E., Mahoney, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M.,  
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,  
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,  
Nguyen, N., Nickerson, E., Nwokweno, S., Ogun, M., Okunnu, G.,  
Oregunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,  
Peters, L., Pickens, R., Prins, E., Pu, L.L., Quiles, M., Ren, Y.,  
Rivers, M., Rojas, A., Rojokokan, I., Rolfe, M., Ruiz, S., Savary, G.,  
Scheiner, S., Scott, G., Shen, H., Shostakov, N., Sisson, I.,  
Sodergren, E., Sonalke, T., Sparks, A., Stanley, H., Stone, H.,  
Sutton, A., Swalek, A., Tabor, P., Tameris, A., Tameris, K., Tang, H.,  
Tansey, J., Taylor, C., Taylor, T., Telirod, B., Thomas, N., Thomas, S.,  
Usmani, K., Vasquez, L., Vera, V., Villalob, D., Vinson, R., Wang, Q.,  
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,  
Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K.,  
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,  
Weinstock, G. and Gibbs, R.

Direct Submission  
Unpublished  
2 (bases 1 to 139754)  
Worley, K.C.

Direct Submission  
Submitted (15-APR-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA.  
3 (bases 1 to 139754)  
Worley, K.C.

Direct Submission  
Submitted (14-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA.  
On Jul 14, 2002 this sequence version replaced gi:20149387.

\*\*\*\*\* Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
Project Information  
Center project name: GVL  
Center clone name: CH230-501B21  
\*\*\*\*\* Summary Statistics  
Sequencing vector: Plasmid  
Chemistry: Dye-terminator Big Dye 100% of reads



Assembly program: Phrap; version 0.990329  
Consensus quality: 116736 bases at least Q40  
Consensus quality: 120062 bases at least Q30  
Consensus quality: 122808 bases at least Q20  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.ngsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.ngsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 30 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
1 1055: contig of 1055 bp in length  
\* 1056 1155: gap of unknown length  
\* 1156 2177: contig of 1022 bp in length  
\* 2178 2277: gap of unknown length  
\* 2278 3508: contig of 1231 bp in length  
\* 3509 3608: gap of unknown length  
\* 3609 5007: contig of 1399 bp in length  
\* 5008 5107: gap of unknown length  
\* 5108 6519: contig of 1412 bp in length  
\* 6520 6619: gap of unknown length  
\* 6620 8203: contig of 1584 bp in length  
\* 8204 8303: gap of unknown length  
\* 8304 9934: contig of 1631 bp in length  
\* 9935 10034: gap of unknown length  
\* 10035 12196: contig of 2162 bp in length  
\* 12197 12296: gap of unknown length  
\* 12297 13460: contig of 1164 bp in length  
\* 13461 13560: gap of unknown length  
\* 13561 16243: contig of 2683 bp in length  
\* 16244 16343: gap of unknown length  
\* 16344 17544: contig of 1201 bp in length  
\* 17545 17644: gap of unknown length  
\* 17645 19497: contig of 1853 bp in length  
\* 19498 19597: gap of unknown length  
\* 19598 21911: contig of 2314 bp in length  
\* 21912 22011: gap of unknown length  
\* 22012 24306: contig of 2295 bp in length  
\* 24307 24406: gap of unknown length  
\* 24407 27677: contig of 3271 bp in length  
\* 27678 27777: gap of unknown length  
\* 27778 31649: contig of 3872 bp in length  
\* 31650 31749: gap of unknown length  
\* 31750 35716: contig of 3967 bp in length  
\* 35717 35816: gap of unknown length  
\* 35817 39569: contig of 3753 bp in length  
\* 39570 39669: gap of unknown length  
\* 39671 42672: contig of 3003 bp in length  
\* 42673 42772: gap of unknown length  
\* 42773 47526: contig of 4754 bp in length  
\* 47527 47626: gap of unknown length  
\* 47627 54170: contig of 6544 bp in length  
\* 54171 54270: gap of unknown length  
\* 54271 58745: contig of 4475 bp in length  
\* 58746 58845: gap of unknown length  
\* 58846 67020: contig of 8175 bp in length  
\* 67021 67120: gap of unknown length  
\* 67121 75578: contig of 8458 bp in length  
\* 75579 75678: gap of unknown length  
\* 75679 81926: contig of 6248 bp in length  
\* 81927 82026: gap of unknown length  
\* 82027 88275: contig of 6249 bp in length  
\* 88276 88375: gap of unknown length  
\* 88376 94434: contig of 6053 bp in length  
\* 94435 94534: gap of unknown length  
\* 94535 103112: contig of 8578 bp in length  
\* 103113 103212: gap of unknown length  
\* 103213 110404: contig of 7192 bp in length  
\* 110405 110504: gap of unknown length

FEATURES  
source  
1. 139754  
/organism="Rattus norvegicus"  
/db\_xref="taxon:10116"  
/clone="CH230-501B21"  
BASE COUNT 39024 a 27442 c 27265 g 36974 t 9049 others  
ORIGIN  
Alignment Scores:  
Pred. No.: 3 47e+03 Length: 139754  
Score: 55.00 Matches: 11  
Percent Similarity: 93.33% Conservative: 3  
Best Local Similarity: 73.33% Mismatches: 1  
Query Match: 73.33% Indels: 0  
DB: 2 Gaps: 0  
US-10-075-846-4\_COPY\_192\_207 (1-16) x AC118327 (1-139754)  
QY 2 SerSerIleuCySerProLeuProSerLeuSerIleuSerVal 16  
:::|||||  
Db 96722 AATTCAATCTTGTCGCCACCCGCGACTCTCTCTCTCTC 96766  
RESULT 14  
AC095184 164772 bp DNA linear HTG 11-JUL-2002  
AC095184 Rattus norvegicus clone CH230-9D7, \*\*\* SEQUENCING IN PROGRESS \*\*\*  
DEFINITION  
85 unordered pieces.  
AC095184  
AC095184.3 GI:21722570  
VERSION  
HTG: HTGS\_PHASE1.  
KEYWORDS  
Norway rat.  
SOURCE  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE  
1 (bases 1 to 164772)  
AUTHORS  
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,  
Alstbrooks,S.L., Amarantunga,H.C., Are,J.R., Ayala,M., Banks,T.,  
Barberia,J., Benton,J., Bimge,K., Blankenburg,K., Bonini,D.,  
Bouch,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,  
Bunay,C., Burch,P., Burkett,C., Butrell,K.L., Byrd,N.C.,  
Carton,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,  
Cleveland,L.C.D., Cox,C., Coyle,M.D., Dathore,S.R., David,R.,  
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,  
Douthwaite,K.U., Draper,H., Dugan-Rocha,S., Durkin,K.J.,  
Einhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
Gabriel,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
Gorelli,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,  
Harris,C., Harris,K., Hart,M., Hayak,P., Hawes,A., Hernandez,J.,  
Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,  
Hornst,E., Howard,S., Huber,J., Huiyk,S., Hume,J., Jackson,L.E.,  
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,  
Kratovich,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louisedge,H.,  
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,  
Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,  
Miner,G., Miner,Z., Mitchell,T., Mohabath,K., Morgan,M., Morris,S.,  
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,  
Nguyen,N., Nickerson,E., Nwokwkw,S., Oguh,M., Okunwo,G.,  
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,D., Perez,L.,  
Peters,L., Pickens,R., Pitmus,E., Pu,L.L., Quiles,M., Ren,Y.,  
Rivers,M., Rojas,A., Rojokokan,I., Rolfe,M., Ruiz,S., Saverly,G.,  
Scherer,S., Scott,G., Shen,H., Shoshari,N., Sisson,I.,  
Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,  
Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
Tansley,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,

Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wallington, S.,  
Williams, G., Williamson, A., Wleczky, R., Wooden, S., Morley, K.,  
Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D.,  
Meinstock, G. and Gibbs, R.  
Direct Submission  
Unpublished  
JOURNAL  
TITLE  
2 (bases 1 to 164772)  
REFERENCE  
AUTHORS  
Morley, K.C.  
JOURNAL  
TITLE  
Submitted (16-SEP-2001) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 164772)  
Morley, K.C.  
Direct Submission  
Submitted (11-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Jul 10, 2002 this sequence version replaced g1:17942274.  
----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GDAM  
Center clone name: CH230-9D7  
----- Summary Statistics  
Sequencing vector: Plasmid;  
Chemistry: Dye-terminator Big Dye 1000 of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 79000 bases at least Q40  
Consensus quality: 85421 bases at least Q30  
Consensus quality: 90756 bases at least Q20  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 85 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\*  
\* 1 1297: contig of 1297 bp in length  
\* 1398 1397: gap of unknown length  
\* 2680 2679: contig of 1282 bp in length  
\* 2780 2779: gap of unknown length  
\* 3803 3802: contig of 1023 bp in length  
\* 3903 3902: gap of unknown length  
\* 4952 4951: contig of 1049 bp in length  
\* 5052 5051: gap of unknown length  
\* 6300 6299: contig of 1248 bp in length  
\* 6400 6399: gap of unknown length  
\* 7432 7431: contig of 1032 bp in length  
\* 7532 7531: gap of unknown length  
\* 8549 8548: contig of 1017 bp in length  
\* 8649 8648: gap of unknown length  
\* 9775 9774: contig of 1128 bp in length  
\* 9876 9875: gap of unknown length  
\* 10877 10876: contig of 1001 bp in length  
\* 10978 10977: gap of unknown length  
\* 12068 12067: contig of 1090 bp in length  
\* 12167 12166: gap of unknown length  
\* 13238 13237: contig of 1072 bp in length  
\* 13340 13339: gap of unknown length  
\* 14778 14777: contig of 1439 bp in length  
\* 14879 14878: gap of unknown length  
\* 16415 16414: contig of 1537 bp in length  
\* 16516 16515: gap of unknown length  
\* 17592 17591: contig of 1077 bp in length  
\* 17593 17592: gap of unknown length  
-----  
\* 17693 18970: contig of 1278 bp in length  
\* 18971 19070: gap of unknown length  
\* 19071 20403: contig of 1333 bp in length  
\* 20404 20503: gap of unknown length  
\* 20504 21651: contig of 1148 bp in length  
\* 21652 21751: gap of unknown length  
\* 21752 22887: contig of 1136 bp in length  
\* 22888 22987: gap of unknown length  
\* 22988 24446: contig of 1459 bp in length  
\* 24447 24546: gap of unknown length  
\* 24547 25565: contig of 1019 bp in length  
\* 25566 25665: gap of unknown length  
\* 25666 27036: contig of 1371 bp in length  
\* 27037 27137: gap of unknown length  
\* 27137 28479: contig of 1342 bp in length  
\* 28479 28578: gap of unknown length  
\* 28579 29664: contig of 1086 bp in length  
\* 29665 31068: gap of unknown length  
\* 31069 31168: contig of 1304 bp in length  
\* 31169 32406: gap of unknown length  
\* 32407 32506: contig of 1238 bp in length  
\* 32507 34145: gap of unknown length  
\* 34146 34245: contig of 1639 bp in length  
\* 34246 35811: gap of unknown length  
\* 35812 35911: contig of 1566 bp in length  
\* 35912 37018: gap of unknown length  
\* 37019 37118: contig of 1107 bp in length  
\* 37119 38624: gap of unknown length  
\* 38625 38724: contig of 1506 bp in length  
\* 38725 40409: gap of unknown length  
\* 40410 40509: contig of 1685 bp in length  
\* 40510 42002: gap of unknown length  
\* 42003 42102: contig of 1493 bp in length  
\* 42103 43186: gap of unknown length  
\* 43187 43259: contig of 1056 bp in length  
\* 43260 44262: gap of unknown length  
\* 44263 44362: contig of 1004 bp in length  
\* 44363 45568: gap of unknown length  
\* 45569 47129: contig of 1206 bp in length  
\* 47130 47229: gap of unknown length  
\* 47230 48940: contig of 1461 bp in length  
\* 48941 49040: gap of unknown length  
\* 49041 50571: contig of 1711 bp in length  
\* 50572 50671: gap of unknown length  
\* 50672 52207: contig of 1536 bp in length  
\* 52208 52307: gap of unknown length  
\* 52308 53449: contig of 1142 bp in length  
\* 53450 53549: gap of unknown length  
\* 53550 54869: contig of 1320 bp in length  
\* 54870 54969: gap of unknown length  
\* 54970 56026: contig of 1057 bp in length  
\* 56027 56126: gap of unknown length  
\* 56127 58611: contig of 2485 bp in length  
\* 58612 58712: gap of unknown length  
\* 58713 60042: contig of 1331 bp in length  
\* 60043 60142: gap of unknown length  
\* 60143 61451: contig of 1309 bp in length  
\* 61452 61551: gap of unknown length  
\* 61552 63614: contig of 2063 bp in length  
\* 63615 65157: gap of unknown length  
\* 65158 65257: contig of 1443 bp in length  
\* 65258 66957: gap of unknown length  
\* 66958 67057: contig of 1700 bp in length  
\* 67058 69461: gap of unknown length  
\* 69462 69561: contig of 2404 bp in length  
\* 69562 70953: gap of unknown length  
\* 70954 71053: contig of 1392 bp in length  
\* 71054 72992: gap of unknown length  
\* 72993 73092: contig of 1939 bp in length  
\* 73093 75634: gap of unknown length  
\* 75634: contig of 2542 bp in length



```

* 31993 37290: contig of 5298 bp in length
* 37291 37390: gap of 100 bp
* 37391 45486: contig of 8096 bp in length
* 45487 45586: gap of 100 bp
* 45587 53572: contig of 7986 bp in length
* 53573 53672: gap of 100 bp
* 53673 66319: contig of 12647 bp in length
* 66320 66419: gap of 100 bp
* 66420 80154: contig of 13735 bp in length
* 80155 80254: gap of 100 bp
* 80255 94472: contig of 14218 bp in length
* 94473 94572: gap of 100 bp
* 94573 120577: contig of 26005 bp in length
* 120578 120677: gap of 100 bp
* 120678 147069: contig of 26392 bp in length
* 147070 147169: gap of 100 bp
* 147170 174354: contig of 27185 bp in length.

```

# FEATURES

```

source
1. 174354
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="RP23-232G11"
/clone_lib="RPC1-23 Female Mouse BMC"
1. 101
/note="assembly-fragment"
clone_end:SP6
vector_side:left"
202. 1145
/note="assembly-fragment"
1246. 2270
/note="assembly-fragment"
2371. 4066
/note="assembly-fragment"
4167. 5725
/note="assembly-fragment"
5826. 8038
/note="assembly-fragment"
8139. 11089
/note="assembly-fragment"
11190. 13192
/note="assembly-fragment"
13293. 16784
/note="assembly-fragment"
16885. 20475
/note="assembly-fragment"
20576. 27499
/note="assembly-fragment"
27600. 31892
/note="assembly-fragment"
31993. 37290
/note="assembly-fragment"
37391. 45486
/note="assembly-fragment"
45587. 53572
/note="assembly-fragment"
53673. 66319
/note="assembly-fragment"
66420. 80154
/note="assembly-fragment"
80255. 94472
/note="assembly-fragment"
94573. 120577
/note="assembly-fragment"
120678. 147069
/note="assembly-fragment"
147170. 174354
/note="assembly-fragment"
clone_end:T7

```

BASE COUNT 46808 a 39544 c 39585 g 46362 t 2055 others  
 ORIGIN

Alignment Scores:

```

Pred. No.: 4.32e+03 Length: 174354
Score: 55.00 Matches: 11
Percent Similarity: 80.00% Conservative: 1
Best Local Similarity: 73.33% Mismatches: 3
Query Match: 73.33% Indels: 0
DB: 2 Gaps: 0

```

US-10-075-846-4\_COPY\_192\_207 (1-16) x AC108830 (1-174354)

QY 2 SerSeriLeuCySerProLeuProSerLeuSerLeuSerVal 16

DB 21798 AGTGAGATTCTGTGGCAATCCCTTCGCCCTTAAGTTACAAAGTC 21842

Search completed: July 1, 2003, 00:18:12  
 Job time : 249.866 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: June 30, 2003, 21:13:09 ; Search time 12.8143 Seconds  
(without alignments)  
2811.853 Million cell updates/sec

Title: US-10-075-846-4\_COPY\_192\_207

Perfect score: 75

Sequence: 1 SSSILCSPLSLSLSV 16

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n model -DEV=xlp  
-O=/cgn2.1/USPRO\_pool/US10075846/runat\_25062003\_163647\_5141/app\_query.fasta\_1.782  
-DB=N.Geneseq\_101002 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR\_SCORE=pcr -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pcr -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10075846.ecgn.1.1.511 -runat\_25062003\_163647\_5141 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOC  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FCAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N.Geneseq\_101002.\*  
1: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1980.DAT:\*  
2: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1981.DAT:\*  
3: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1982.DAT:\*  
4: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1983.DAT:\*  
5: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1984.DAT:\*  
6: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1985.DAT:\*  
7: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1986.DAT:\*  
8: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1987.DAT:\*  
9: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1988.DAT:\*  
10: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1989.DAT:\*  
11: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1990.DAT:\*  
12: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1991.DAT:\*  
13: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1992.DAT:\*  
14: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1993.DAT:\*  
15: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1994.DAT:\*  
16: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1995.DAT:\*  
17: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1996.DAT:\*  
18: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1997.DAT:\*  
19: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1998.DAT:\*  
20: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1999.DAT:\*  
21: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2000.DAT:\*  
22: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*  
23: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*  
24: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the total being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	68	90.7	39796	21 AAC61681	Nucleotide sequenc
2	53	70.7	3294	14 AAQ48939	Steroid hormone re
3	53	70.7	3301	20 AAX36523	Human PPAR-delta c
4	53	70.7	3301	24 ABL65244	Lung cancer relate
5	52	69.3	90	22 ABA75659	Human foetal liver
6	52	69.3	90	22 AAK24269	Human brain expres
7	52	69.3	90	22 AAI27367	Probe #17300 for g
8	52	69.3	597	22 ABA63211	Human foetal liver
9	52	69.3	597	22 AAK1674	Human brain expres
10	52	69.3	597	22 AAI18235	Probe #8168 for ge
11	52	69.3	3599	22 AAD17468	Human transporter
12	51	68.0	20453	23 ABL03244	Drosophila melanog
13	50	66.7	418	23 ABV16725	Human prostate exp
14	50	66.7	471	23 ABV46523	Human prostate exp
15	50	66.7	1689	20 AAZ23024	Rat kd312 polypept
16	50	66.7	3079	20 AAZ23025	Rat kd312 genomic
17	49	65.3	1082	23 AAS74624	DNA encoding novel
18	49	65.3	2667	21 AAC45706	Arabidopsis thalia
19	49	65.3	3171	23 AAS74569	DNA encoding novel
20	49	65.3	3171	23 AAS75059	DNA encoding novel
21	49	65.3	112460	24 ABR83567	Human cDNA differe
22	49	65.3	249487	24 ABR85733	Mouse genomic regi
23	48	64.0	288	22 AAL16141	Human breast cance
24	48	64.0	289	22 AAL24984	Human breast cance
25	48	64.0	576	24 ABR73574	Bacillus licheniflo
26	48	64.0	668	22 AAK68950	Human immune/haema
27	48	64.0	698	22 AAI94991	Human neuroblastom
28	48	64.0	730	22 AAI95568	Human neuroblastom
29	48	64.0	827	21 AAC47631	Arabidopsis thalia
30	48	64.0	3271	19 AAUV0060	Human G-protein co
31	48	64.0	4010	20 AAV72212	Rat borta ranBC pr
32	48	64.0	5412	20 AAX21356	Human BA12 gene.
33	48	64.0	10468	22 AAS30509	DNA encoding novel
34	48	64.0	10468	22 AAS30512	DNA encoding novel
35	48	64.0	10468	22 AAL06288	Human reproductive
36	48	64.0	10468	22 AAL06291	Human reproductive
37	47.5	63.3	19408	22 AAS42003	Genomic sequence #
38	47.5	63.3	19408	22 AAK87230	Human immune/haema
39	47.5	63.3	19408	22 AAK90644	Human digestive sy
40	47	62.7	144	23 ABL26393	Drosophila melanog
41	47	62.7	255	22 AAL23844	Human breast cance
42	47	62.7	641	24 ABR82059	Novel floral meris
43	47	62.7	641	24 ABR82060	Novel floral meris
44	47	62.7	666	24 ABR82062	Novel floral meris
45	47	62.7	731	24 ABR82061	Novel floral meris

## ALIGNMENTS

RESULT 1  
AAC61681  
ID AAC61681 standard: DNA: 39796 BP.

AC AAC61681:  
19-FEB-2001 (first entry)

DE Nucleotide sequence of the human ataxia gene.

KW Human: ataxia: gene therapy: ss.

OS Homo sapiens.

FT Key Location/Qualifiers  
FT exon 29850..29921  
FT intron 29922..33025

```
FT FT exon /tag- b
FT FT 33026..33155
FT FT /tag- c
FT FT 33156..33444
FT FT intron /tag- d
FT FT 33445..33514
FT FT exon /tag- e
FT FT intron /tag- f
FT FT 33752..33975
FT FT exon /tag- g
FT FT intron /tag- h
FT FT 33976..34114
FT FT exon /tag- i
FT FT 34115..34195
FT FT intron /tag- j
FT FT 34196..35759
FT FT exon /tag- k
FT FT 35760..35901
FT FT intron /tag- l
FT FT 35902..38781
FT FT exon /tag- m
FT FT 38782..38996
FT FT /tag- m
XX PN WO200058461-A1.
XX PD 05-OCT-2000.
XX PF 23-MAR-2000; 2000WO-EP02600.
XX PR 26-MAR-1999; 99EP-0106343.
XX PA (RAPP/) RAPPOLD-HOERBRAND G.
XX PI Rappold-Hoerbrand G;
XX DR MPI; 2000-656166/63.
XX PT Novel nucleic acid sequence encoding human ataxia protein for screening
XX PT compounds useful for treating disorders relating to mutations in ataxia
XX PT gene -
XX PS Claim 6; Page 22-44; 47pp; English.
XX CC The present sequence represents the human ataxia gene. The ataxia
XX CC protein and polynucleotides are useful for diagnosing and treating
XX CC disorders related to ataxia. Ataxia gene sequences are useful in
XX CC gene therapy, and as diagnostic tools or reagents for identifying and
XX CC characterizing genetic defect involved in the disorders and diseases
XX CC related to ataxia.
XX SQ Sequence 39796 BP; 10630 A; 9337 C; 9355 G; 10474 T; 0 other;

Alignment Scores:
Pred. No.: 30.5 Length: 39796
Score: 68.00 Matches: 14
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 93.33% Mismatches: 0
Query Match: 90.67% Indels: 0
DB: 21 Gaps: 0

US-10-075-846-4_COPY_192_207 (1-16) x AAC61681 (1-39796)
OY 2 SerSerIleLeuCySerProLeuProSerLeuSerLeuSerVal 16
Db 35719 GCATCATACTCTGCAGCCCTCTGCCATCTCTGTCTTCACTTCACTT 35763

RESULT 2
AAQ48939
ID AAQ48939 standard; DNA; 3294 BP.
XX AC AAQ48939;
XX XX
```

```
DT 13-APR-1994 (first entry)
XX XX
DE Steroid hormone receptor (NUCI) gene.
XX XX
KW Steroid; hormone; receptor; osteosarcoma; superfamily; ss.
XX XX
OS Homo sapiens.
XX XX
FH Key Location/Qualifiers
FT CDS 338..1663
FT FT /tag- a
FT FT /product= Steroid hormone receptor.
XX PN GB2265376-A.
XX PD 29-SEP-1993.
XX PF 23-MAR-1993; 93GB-0006043.
XX PR 24-MAR-1992; 92US-0857055.
XX PA (MERI ) MERCK & CO INC.
XX XX
XX PI Rodan GA, Rutledge SJ, Schmidt A, Vogel RL;
XX DR MPI; 1993-305586/39.
XX DR P-PSDB; AAR41875.
XX PT New human steroid hormone receptor NUCI - used to assay and
XX PT evaluate ligands binding to the receptor.
XX PS Claim 7; Page 6-8; 62pp; English.
XX CC The steroid hormone receptor (designated NUCI) is used in assays to
XX CC identity and evaluate chemical entities that bind to it.
XX SQ Sequence 3294 BP; 705 A; 1014 C; 879 G; 696 T; 0 other;

Alignment Scores:
Pred. No.: 309 Length: 3294
Score: 53.00 Matches: 12
Percent Similarity: 92.86% Conservative: 1
Best Local Similarity: 85.71% Mismatches: 1
Query Match: 70.67% Indels: 0
DB: 14 Gaps: 0

US-10-075-846-4_COPY_192_207 (1-16) x AAQ48939 (1-3294)
OY 2 SerSerIleLeuCySerProLeuProSerLeuSerLeuSer 15
Db 1837 TCTTCCTCTCTCCCTTCCCTTCCCTTCCCTTCTCTCTCC 1878

RESULT 3
AAK36523
ID AAK36523 standard; cDNA; 3301 BP.
XX AC AAK36523;
XX XX
DT 07-JUL-1999 (first entry)
XX DE Human PPAR-delta coding sequence.
XX DE
XX DE Nuclear receptor agonist; antagonist; identification; PPAR;
XX DE peroxisome proliferator activated receptor; ss.
XX OS Homo sapiens.
XX PN WO9918124-A1.
XX PD 15-APR-1999.
XX PF 06-OCT-1998; 98WO-US21049.
XX XX
```





OY 2 SerSerileuCySerProleuProSerleuSerleuSer 15  
|||||:|||||  
DB 1837 TCTTCCTTCCTCCCTTCCTCCCTTCCTTCCTCC 1878

## RESULT 5

ABAT75659/c  
ID ABA75659 standard; DNA; 90 BP.

AC ABA75659;

XX 01-FEB-2002 (first entry)

DE Human foetal liver single exon nucleic acid probe #23964.

KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.

OS Homo sapiens.

PN WO200157277-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00669.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-063366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

PT WPI: 2001-483447/52.

PS Claim 4; SEQ ID NO 23964; 639pp + sequence listing; English.

CC The invention relates to a single exon nucleic acid probe for

CC measuring human gene expression in a sample derived from human foetal

CC liver. The single exon nucleic acid probes may be used for predicting,

CC measuring and displaying gene expression in samples derived from human

CC foetal liver. The present sequence is a single exon nucleic acid

CC probe of the invention.

CC Note: The sequence data for this patent did not form part of the

CC printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 90 BP; 29 A; 16 C; 27 G; 18 T; 0 other;

XX Alignment Scores:

XX Pred. No.: 5.98 Length: 90

XX Score: 52.00 Matches: 10

XX Percent Similarity: 92.31% Conservative: 2

XX Best Local Similarity: 76.92% Mismatches: 1

XX Query Match: 69.33% Indels: 0

XX DB: 22 Gaps: 0

US-10-075-846-4\_COPY\_192\_207 (1-16) x ABA75659 (1-90)

OY 3 SerileuCySerProleuProSerleuSerleuSer 15

DB 81 TCGTATCTGCTCTCTCCCTTCCTGCTCTCTCC 43

## RESULT 6

AAK24269/c  
ID AAK24269 standard; DNA; 90 BP.

AC AAK24269;

XX 05-NOV-2001 (first entry)

DE Human brain expressed single exon probe SEQ ID NO: 24260.

KW Human; brain expressed exon; gene expression analysis; probe;

KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;

KW epilepsy; cancer; ss.

OS Homo sapiens.

PN WO200157275-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00667.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-063366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

PT WPI: 2001-483446/52.

PS Example 4; SEQ ID NO: 24260; 650pp + Sequence Listing; English.

CC The present invention provides a number of single exon nucleic acid

CC probes which are derived from genomic sequences expressed in the human

CC brain. They can be used to measure gene expression in brain cell samples,

CC which may enable the diagnosis and improved treatment of nervous system

CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,

CC epilepsy and cancers. The present sequence is one of the probes of the

XX SQ Sequence 90 BP; 29 A; 16 C; 27 G; 18 T; 0 other;

XX Alignment Scores:

XX Pred. No.: 5.98 Length: 90

XX Score: 52.00 Matches: 10

XX Percent Similarity: 92.31% Conservative: 2

XX Best Local Similarity: 76.92% Mismatches: 1

XX Query Match: 69.33% Indels: 0

XX DB: 22 Gaps: 0

US-10-075-846-4\_COPY\_192\_207 (1-16) x AAK24269 (1-90)

OY 3 SerileuCySerProleuProSerleuSerleuSer 15

DB 81 TCGTATCTGCTCTCTCCCTTCCTGCTCTCTCC 43

AAI27367/c

ID AAI27367 standard; DNA; 90 BP.

AC AAI27367;

XX 12-OCT-2001 (first entry)

DE Probe #17300 for gene expression analysis in human cervical cell sample.

KW Probe; human; microarray; gene expression; cervical epithelial cell;

KW cervical cancer; ss.

```
XX OS Homo sapiens.
XX PN WO200157278-A2.
XX XX
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US00670.
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-488901/53.
XX PT Human genome-derived single exon nucleic acid probes useful for
XX PT analyzing gene expression in human cervical epithelial cells -
XX PS Claim 25; SEQ ID NO 17300; 487pp; English.
XX CC The present invention relates to human single exon nucleic acid probes
XX CC (SENP). The present sequence is one such probe. The SENPs are derived
XX CC from human HeLa cells. The SENPs can be used to produce a single exon
XX CC microarray, which can be used for measuring human gene expression in a
XX CC sample derived from human cervical epithelial cells. By measuring gene
XX CC expression, the probes are therefore useful in grading and/or staging
XX CC of diseases of the cervix, notably cervical cancer.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 90 BP; 29 A; 16 C; 27 G; 18 T; 0 other;

Alignment Scores:
Pred. No.: 5.98 Length: 90
Score: 52.00 Matches: 10
Percent Similarity: 92.31% Conservative: 2
Best Local Similarity: 76.92% Mismatches: 1
Query Match: 69.33% Indels: 0
DB: 22 Gaps: 0

US-10-075-846-4_COPY_192_207 (1-16) x AAI27367 (1-90)
OY 3 SerileucySserProleuproSerleuSerleuSer 15
ID ABA63211/C
XX ABA63211 standard; DNA; 597 BP.
XX AC ABA63211:
XX DT 01-FEB-2002 (first entry)
XX DE Human foetal liver single exon nucleic acid probe #11516.
XX DE Human foetal liver; gene expression; single exon nucleic acid probe: ss.
XX KW Human; foetal liver; gene expression; single exon nucleic acid probe: ss.
XX OS Homo sapiens.
XX PN WO200157277-A2.
XX PD 09-AUG-2001.
XX XX
```

```
PF 30-JAN-2001; 2001WO-US00669.
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-483447/52.
XX PT Human genome-derived single exon nucleic acid probes useful for
XX PT analyzing gene expression in human fetal liver -
XX PS Claim 1; SEQ ID NO 11516; 639pp + sequence listing; English.
XX CC The invention relates to a single exon nucleic acid probe for
XX CC measuring human gene expression in a sample derived from human foetal
XX CC liver. The single exon nucleic acid probes may be used for predicting,
XX CC measuring and displaying gene expression in samples derived from human
XX CC fetal liver. The present sequence is a single exon nucleic acid
XX CC probe of the invention.
XX CC Note: The sequence data for this patent did not form part of the
XX CC printed specification, but was obtained in electronic format directly
XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 597 BP; 186 A; 117 C; 116 G; 178 T; 0 other;

Alignment Scores:
Pred. No.: 57.2 Length: 597
Score: 52.00 Matches: 10
Percent Similarity: 92.31% Conservative: 2
Best Local Similarity: 76.92% Mismatches: 1
Query Match: 69.33% Indels: 0
DB: 22 Gaps: 0

US-10-075-846-4_COPY_192_207 (1-16) x ABA63211 (1-597)
OY 3 SerileucySserProleuproSerleuSerleuSer 15
ID AAK11674/C
XX AAK11674 standard; DNA; 597 BP.
XX AC AAK11674:
XX DT 05-NOV-2001 (first entry)
XX DE Human brain expressed single exon probe SEQ ID NO: 11665.
XX DE Human brain expressed exon; gene expression analysis; probe:
XX KW Human; brain expressed exon; gene expression analysis; probe:
XX KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
XX KW epilepsy; cancer; ss.
XX OS Homo sapiens.
XX PN WO200157275-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US00667.
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX XX
```

```

PA      (MOLE-) MOLECULAR DYNAMICS INC.
XX      Penn SG, Hanzel DK, Chen W, Rank DR;
XX      WPI; 2001-483446/52.
DR      Single exon nucleic acid probes for analyzing gene expression in human
        brains -
PS      Example 4; SEQ ID NO: 11665; 650bp + Sequence Listing; English.
CC      The present invention provides a number of single exon nucleic acid
CC      probes which are derived from genomic sequences expressed in the human
CC      brain. They can be used to measure gene expression in brain cell samples
CC      which may enable the diagnosis and improved treatment of nervous system
CC      diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC      epilepsy and cancers. The present sequence is one of the probes of the
CC      invention.
XX      SQ      Sequence 597 BP; 186 A; 117 C; 116 G; 178 T; 0 other;
        Alignment Scores:
        Pred. No.:          57.2           Length:          597
        Score:              52.00          Matches:          10
        Percent Similarity: 92.31%         Conservative:     2
        Best Local Similarity: 76.92%       Mismatches:      1
        Query Match:        69.33%         Indels:          0
        DB:                  22            gaps:          0
US-10-075-846-4_COPY_192_207 (1-16) x AAk11674 (1-597)
OY      3 SerIleLeuCysSerProIeuProSerIeuSerIeuSer 15
        |||:::|||||||
DB      443 TCCTACTTCGTCGTCTCCTCTCCCTCGTCTCTCTCC 405
RESULT 10
AA118235/C
ID      AA118235 standard; DNA; 597 BP.
XX      AC      AA118235;
XX      DT      12-Oct-2001 (first entry)
XX      DE      Probe #8168 for gene expression analysis in human cervical cell sample.
XX      EE      Probe: human; microarray; gene expression; cervical epithelial cell;
XX      KW      cervical cancer; ss.
XX      OS      Homo sapiens.
XX      PN      MO200157278-A2.
XX      PD      09-AUG-2001.
XX      PE      30-JAN-2001; 2001WO-US00670.
XX      PR      04-FEB-2000; 2000US-0180312.
XX      PR      26-MAY-2000; 2000US-0207456.
XX      PR      30-JUN-2000; 2000US-0608408.
XX      PR      03-AUG-2000; 2000US-0632366.
XX      PR      21-SEP-2000; 2000US-0234687.
XX      PR      27-SEP-2000; 2000US-0236359.
XX      PR      04-OCT-2000; 2000GB-0024263.
XX      PA      (MOLE-) MOLECULAR DYNAMICS INC.
XX      PI      Penn SG, Hanzel DK, Chen W, Rank DR;
XX      WI      WPI; 2001-488901/53.
XX      WI      WPI; 2001-488901/53.

```

XX	Human genome-derived single exon nucleic acid probes useful for
PT	analyzing gene expression in human cervical epithelial cells -
XX	
PS	Claim 25; SEQ ID No 8168; 487bp; English.
XX	
CC	The present invention relates to human single exon nucleic acid probes
CC	(SENP). The present sequence is one such probe. The SENPs are derived
CC	from human HeLa cells. The SNPs can be used to produce a single exon
CC	microarray, which can be used for measuring human gene expression in a
CC	sample derived from human cervical epithelial cells. By measuring gene
CC	expression, the probes are therefore useful in grading and/or staging
CC	of diseases of the cervix, notably cervical cancer.
CC	Note: The sequence data for this patent did not form part of the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at ftp.wipo.int/pub/published_pct-sequences.
XX	
SQ	Sequence 597 BP; 186 A; 117 C; 116 G; 178 T; 0 other;
	Alignment Scores:
	Pred. No.: 57.2 Length: 597
	Score: 52.00 Matches: 10
	Percent Similarity: 92.31% Conservative: 2
	Best Local Similarity: 76.92% Mismatches: 1
	Query Match: 69.33% Indels: 0
	DB: Gaps: 0
	US-10-075-846-4_COPY_192_207 (1-16) x AAI18235 (1-597)
Oy	3 SerLeucenySerProLeuProSerLeuSerLeuSer 15
	:
Db	443 TCGCTATCTTGCCTCTCCCTCTCCCTGCTCTCTCC 405
RESULT 11	
	AAI17468/c
ID	AAI17468 standard; cDNA; 3599 BP.
XX	
AC	AAI17468;
XX	
DT	10-DEC-2001 (first entry)
XX	
DE	Human transporter and ion channel-1 (TRICH-1) cDNA.
XX	
KW	Human; transporter and ion channel; TRICH-1; therapy; akinesia; cardiant;
KW	neurological disorder; immune disorder; allergy; nootropic; dementia;
KW	AIDS; Acquired Immune Deficiency Syndrome; amyotrophic lateral sclerosis;
KW	cystic fibrosis; Alzheimer's disease; depression; epilepsy; scleroderma;
KW	cardiomyopathy; muscular dystrophy; Crohn's disease; multiple sclerosis;
KW	rheumatoid arthritis; cancer; anaemia; asthma; trauma; haemodialysis;
KW	gastritis; inflammation; ss.
XX	
OS	Homo sapiens.
XX	
FFH	Location/Qualifiers
FT	CDS 221..1168
FT	/*tag= a
FT	/product= "Human TRICH-1 protein"
XX	
PN	WO200162923-A2.
XX	
PD	30-AUG-2001.
XX	
PE	23-FEB-2001; 2001WO-US05942.
XX	
PR	25-FEB-2000; 2000US-0184866.
PR	02-MAR-2000; 2000US-0187947.
PR	09-MAR-2000; 2000US-0188333.
PR	17-MAR-2000; 2000US-0190230.
PR	24-MAR-2000; 2000US-0192077.
PR	30-MAR-2000; 2000US-0193500.
XX	
PA	(INCYTE-) INCYTE GENOMICS INC.
XX	



XX The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for:  
CC (a) assessing whether a patient is afflicted with prostate cancer;  
CC (b) monitoring the progression of prostate cancer in a patient;  
CC (c) assessing the efficacy of a test compound to inhibit prostate  
CC cancer in a patient;  
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer  
CC in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound;  
CC (g) determining whether prostate cancer has metastasized in a patient;  
CC (h) assessing the aggressiveness or indolence of prostate cancer in a  
CC patient;  
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.

XX Sequence 418 BP; 133 A; 71 C; 103 G; 111 T; 0 other;

Alignment Scores:  
Pred. No.: 75.7 Length: 418  
Score: 50.00 Matches: 9  
Percent Similarity: 92.86% Conservative: 4  
Best Local Similarity: 64.29% Mismatches: 1  
Query Match: 66.67% Indels: 0  
DB: 23 Gaps: 0

US-10-075-846-4\_COPY\_192\_207 (1-16) x ABV16725 (1-418)

QY 1 SerSerSerIleLeuCySerProLeuProSerLeuSerLeu 14

DB 161 TCAGCACCTGTCTGTCTCCCTCCCTACTGTGACGCTT 120

RESULT 14

ABV46523/C  
ID ABV46523 standard; cDNA; 471 BP.

XX AC ABV46523;

XX DT 16-SEP-2002 (first entry)

XX DE Human prostate expression marker cDNA 46514.

XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
XX pharmacogenomic marker; gene; ss.

XX OS Homo sapiens.

XX PN WO200160860-A2.

XX PD 23-AUG-2001.

XX PF 20-FEB-2001; 2001WO-US05171.

XX PR 17-FEB-2000; 2000US-183319P.

XX PR 16-MAR-2000; 2000US-189662P.

XX PR 25-MAY-2000; 2000US-207454P.

XX PR 09-JUN-2000; 2000US-211314P.

XX PR 18-JUL-2000; 2000US-219007P.

XX PR 13-DEC-2000; 2000US-255281P.

XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Schlegel R, Endege WO, Monahan JE;

XX DR WPI; 2001-662795/76.

XX PT Novel isolated nucleic acid molecule associated with cancerous state of  
XX prostate cells and correlating with presence of prostate cancer, useful  
XX for detecting presence of prostate cancer, stage of prostate cancer  
XX Claim 1; Page 9177; 11750pp; English.

CC The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for:  
CC (a) assessing whether a patient is afflicted with prostate cancer;  
CC (b) monitoring the progression of prostate cancer in a patient;  
CC (c) assessing the efficacy of a test compound to inhibit prostate  
CC cancer in a patient;  
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer  
CC in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound;  
CC (g) determining whether prostate cancer has metastasized in a patient;  
CC (h) assessing the aggressiveness or indolence of prostate cancer in a  
CC patient;  
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.

XX Sequence 471 BP; 149 A; 89 C; 115 G; 118 T; 0 other;

Alignment Scores:  
Pred. No.: 87.3 Length: 471  
Score: 50.00 Matches: 9  
Percent Similarity: 92.86% Conservative: 4  
Best Local Similarity: 64.29% Mismatches: 1  
Query Match: 66.67% Indels: 0  
DB: 23 Gaps: 0

US-10-075-846-4\_COPY\_192\_207 (1-16) x ABV46523 (1-471)

QY 1 SerSerSerIleLeuCySerProLeuProSerLeuSerLeu 14

DB 214 TCAGCACCTGTCTGTCTCCCTCCCTACTGTGACGCTT 173

RESULT 15

AA223024  
ID AA223024 standard; cDNA; 1689 BP.

XX AC AA223024;

XX DT 17-JAN-2000 (first entry)

XX DE Rat kd312 polypeptide encoding cDNA.

XX KW kd312 polypeptide; cancer; immunodeficiency disorder; AIDS; stroke;  
XX heart attack; head trauma; neurodegenerative disease; rat;  
XX Parkinson's disease; Alzheimer's disease; ss.

XX OS Rattus sp.

XX PN WO9950288-A2.

XX PD 07-OCT-1999.

XX PF 30-MAR-1999; 99WO-US06993.

XX PR 31-MAR-1998; 98US-0053374.

XX PA (AMGE-) AMGEN INC.

XX PI Yen K;

XX WPI; 1999-601322/51.

XX DR P-PSDB; AAY42694.

XX PT kd312 polypeptides useful for treating diseases and disorders  
XX associated with alterations in cell proliferation and cell death

XX PS Claim 2; Fig 9; 85pp; English.

XX The invention provides nucleic acid molecules encoding human and rat  
XX kd312 polypeptides. The kd312 polypeptides can be expressed by standard  
XX recombinant methodology. The kd312 sequences, and the antibodies against  
XX the proteins may be used to treat or diagnose the presence or progression  
XX of various diseases, e.g. cancer, immunodeficiency disorders (e.g. AIDS),

CC stroke, heart attack, head trauma, and neurodegenerative diseases (e.g. Parkinson's disease and Alzheimer's disease). The present sequence represents the rat kd312 cDNA sequence.

XX  
SO Sequence 1689 BP; 417 A; 462 C; 434 G; 376 T; 0 other;

## Alignment Scores:

Pred. No.:	401	Length:	1689
Score:	50.00	Matches:	10
Percent Similarity:	83.33%	Conservative:	0
Best Local Similarity:	83.33%	Mismatches:	2
Query Match:	66.67%	Indels:	0
DB:	20	Gaps:	0

US-10-075-846-4\_COPY\_192\_207 (1-16) x AA223024 (1-1689)

OY	2	SerSerIleIeuCysSerProIeuProSerIeuSer	13
Db	66	TCCGGCGCCCTCTGCACGCCCTCTACCTCTCTCAGC	101

Search completed: June 30, 2003, 23:13:47  
Job time : 18.8143 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - nucleic search, using frame\_plus.p2n model

Run on: June 30, 2003, 23:02:29 ; Search time 3.40045 Seconds

(without alignments)  
1442.995 Million cell updates/sec

Title: US-10-075-846-4\_COPY\_192\_207

Perfect score: 75  
Sequence: 1 SSILCSPLSLSLSV 16

## Scoring table:

Xgapop 10.0	Xgapext 0.5
Ygapop 10.0	Ygapext 0.5
Fgapop 6.0	Fgapext 7.0
Delop 6.0	Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

## Command line parameters:

```
-MODEL=frame+ p2n.model -DEV=xlp  
-O=/cgn2_1/USPRO_spool/US10075846/runat_25062003.163649.5167/app_query.fasta_1.782  
-DB-Issued-Patents.NA -OEWI-fastcap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=BITS -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pcio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10075846.ecgn.1.1.84.ernat.25062003.163649.5167 -NCPU=6 -ICU=3  
-NO_MMAP -LARGEMEMORY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

## Database :

```
1: Issued.Patents.NA:*  
2: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*  
3: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*  
4: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*  
5: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*  
6: /cgn2_6/ptodata/1/ina/Backfiles1.seq:*
```

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	50	66.7	1689	4	US-09-053-374A-4
2	50	66.7	3079	4	US-09-053-374A-6
3	48	64.0	3271	2	US-08-852-806-1
4	48	64.0	3271	3	US-09-163-666-1
5	46	61.3	724	4	US-09-535-008-40
6	46	61.3	1331	4	US-09-179-221D-3
7	46	61.3	3191	4	US-09-453-702B-70
8	46	61.3	4680	1	US-08-254-358-1
9	46	61.3	4680	1	US-08-475-391-1
10	46	61.3	4680	2	US-08-709-609-1
11	46	61.3	4680	5	PCT-US95-07178-1
12	46	61.3	4910	2	US-08-331-384-2

13	46	61.3	4910	2	US-08-836-087-2	Sequence 2, Appl1
14	46	61.3	4910	4	US-09-246-320-2	Sequence 2, Appl1
15	46	61.3	4910	4	US-09-546-738-2	Sequence 2, Appl1
16	46	61.3	7557	4	US-09-770-315-3	Sequence 3, Appl1
17	46	61.3	8698	4	US-09-770-315-2	Sequence 2, Appl1
18	46	61.3	4403765	4	US-09-103-840A-2	Sequence 2, Appl1
19	46	61.3	4411529	4	US-09-103-840A-1	Sequence 1, Appl1
20	45	60.0	363	4	US-08-580-031A-13	Sequence 13, Appl1
21	45	60.0	1876	4	US-08-840-767-5	Sequence 5, Appl1
22	45	60.0	2234	4	US-09-087-138-7	Sequence 7, Appl1
23	45	60.0	2303	3	US-09-289-376-1	Sequence 1, Appl1
24	45	60.0	2333	2	US-08-732-028-1	Sequence 1, Appl1
25	45	60.0	2333	4	US-09-183-228-1	Sequence 36, Appl1
26	45	60.0	2976	4	US-09-221-017B-13	Sequence 3, Appl1
27	45	60.0	50000	4	US-09-146-053-3	Sequence 43, Appl1
28	44	58.7	517	4	US-09-288-143-43	Sequence 2, Appl1
29	44	58.7	1140	1	US-08-289-653-2	Sequence 2, Appl1
30	44	58.7	2845	1	US-08-289-653-1	Sequence 21, Appl1
31	44	58.7	4609	4	US-09-322-478-21	Sequence 19, Appl1
32	44	58.7	9829	4	US-09-322-478-19	Sequence 1, Appl1
33	44	58.7	11707	4	US-09-136-574A-1	Sequence 17, Appl1
34	44	58.7	12286	4	US-09-322-478-17	Sequence 16, Appl1
35	44	58.7	152331	3	US-09-128-155-16	Sequence 17, Appl1
36	44	58.7	176373	3	US-09-128-155-17	Sequence 9, Appl1
37	43.5	58.0	1906	4	US-09-031-962D-5	Sequence 5, Appl1
38	43.5	58.0	3153	4	US-09-175-928-9	Sequence 1, Appl1
39	43	57.3	2820	5	PCT-US93-11725-1	Sequence 3, Appl1
40	43	57.3	3074	5	PCT-US93-11725-3	Sequence 6, Appl1
41	43	57.3	7015	4	US-09-177-248-6	Sequence 3, Appl1
42	43	57.3	29629	4	US-09-729-995-3	Sequence 11, Appl1
43	42	56.0	185	4	US-09-402-668-11	Sequence 11, Appl1
44	42	56.0	1109	1	US-08-183-214-11	Sequence 89, Appl1
45	42	56.0	1168	4	US-09-484-970B-89	

## ALIGNMENTS

```
RESULT 1  
US-09-053-374A-4  
; Sequence 4, Application US/09053374A  
; Patent No. 6462177  
; GENERAL INFORMATION:  
; APPLICANT: YEN, KWANG-MU  
; TITLE OF INVENTION: MAMMALIAN BLOOD LOSS-INDUCED GENE, KD312  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: AMGEN INC.  
; STREET: ONE AMGEN CENTER DRIVE  
; CITY: THOUSAND OAKS  
; STATE: CA  
; COUNTRY: US  
; ZIP: 91320  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/053,374A  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: COOK, ROBERT R.  
; REGISTRATION NUMBER: 31,602  
; REFERENCE FOR SEQ ID NO: 4:  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1689 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:
```

```

; NAME/KEY: CDS
; LOCATION: 132..971
; US-09-053-374A-4

Alignment Scores:
Pred. No.: 71.8      Length: 1689
Score: 50.00      Matches: 10
Percent Similarity: 83.33%      Conservative: 0
Best Local Similarity: 83.33%      Mismatches: 2
Query Match: 66.67%      Indels: 0
DB: 4      Gaps: 0

US-10-075-846-4_COPY_192_207 (1-16) x US-09-053-374A-4 (1-1689)

QY      2 SerSerIleuCySerProleuProSerleuSer 13
      ||| ||||| ||||| ||||| ||||| |||||
Db      66 TCCGGCGCCCTCTGCAGCCCTCTACTTCTCTCAGC 101

RESULT 2
US-09-053-374A-6
; Sequence 6, Application US/09053374A
; Patent No. 6462177
; GENERAL INFORMATION:
; APPLICANT: YEN, KWANG-MI
; TITLE OF INVENTION: MAMMALIAN BLOOD LOSS-INDUCED GENE, KD312
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AMGEN INC.
; STREET: ONE AMGEN CENTER DRIVE
; CITY: THOUSAND OAKS
; STATE: CA
; COUNTRY: US
; ZIP: 91320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/053,374A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: COOK, ROBERT R.
; REGISTRATION NUMBER: 31,602
; REFERENCE/DOCKET NUMBER: A-514
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3079 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-09-053-374A-6

Alignment Scores:
Pred. No.: 147      Length: 3079
Score: 50.00      Matches: 10
Percent Similarity: 83.33%      Conservative: 0
Best Local Similarity: 83.33%      Mismatches: 2
Query Match: 66.67%      Indels: 0
DB: 4      Gaps: 0

US-10-075-846-4_COPY_192_207 (1-16) x US-09-053-374A-6 (1-3079)

QY      2 SerSerIleuCySerProleuProSerleuSer 13
      ||| ||||| ||||| ||||| ||||| |||||
Db      237 TCCGGCGCCCTCTGCAGCCCTCTACTTCTCTCAGC 272

RESULT 3
US-08-852-806-1
; Sequence 1, Application US/08852806
; Patent No. 5874245

; GENERAL INFORMATION:
; APPLICANT: Shoji Fukusumi
; TITLE OF INVENTION: NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: U.S.A
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/852,806
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/017,915
; FILING DATE: 16 MAY 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Han, William T
; REGISTRATION NUMBER: 34,344
; REFERENCE/DOCKET NUMBER: TAK50002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5219
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3271 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; US-08-852-806-1

Alignment Scores:
Pred. No.: 321      Length: 3271
Score: 48.00      Matches: 9
Percent Similarity: 90.91%      Conservative: 1
Best Local Similarity: 81.82%      Mismatches: 1
Query Match: 64.00%      Indels: 0
DB: 2      Gaps: 0

US-10-075-846-4_COPY_192_207 (1-16) x US-08-852-806-1 (1-3271)

QY      5 LeuCySerProleuProSerleuSerleuSer 15
      |||||:||||| ||||| |||||
Db      1064 CTGTGGGCGCCCTTACCCAGCCTCCAGCTGAGC 1096

RESULT 4
US-09-163-669-1
; Sequence 1, Application US/09163669
; Patent No. 6111076
; GENERAL INFORMATION:
; APPLICANT: FUKUSUMI, SHOJI
; APPLICANT: HINUMA, SHUJI
; APPLICANT: FUJII, RYO
; TITLE OF INVENTION: NOVEL HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTOR (HIBCD07)
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ratner & Prestia
; STREET: P.O. Box 980
```

CITY: Valley Forge  
STATE: PA  
COUNTRY: USA  
ZIP: 19482  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/163.669  
FILING DATE: 30-SEP-1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/852.806  
FILING DATE: 07-MAY-1997  
APPLICATION NUMBER: 60/017.915  
FILING DATE: 16-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Prestia, Paul F.  
REGISTRATION NUMBER: 23.031  
REFERENCE/DOCKET NUMBER: TAK-50002-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-407-0700  
TELEFAX: 610-407-0700  
TELEX: 846169  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3271 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-09-163-669-1

Alignment Scores:  
Pred. No.: 321 Length: 3271  
Score: 48.00 Matches: 9  
Percent Similarity: 90.91% Conservative: 1  
Best Local Similarity: 81.82% Mismatches: 1  
Query Match: 64.00% Indels: 0  
DB: 3 Gaps: 0

US-10-075-846-4\_COPY\_192\_207 (1-16) x US-09-163-669-1 (1-3271)  
QY 5 LeucySerrProleupProSerleuSerleuSer 15  
DB 1064 CTGTGCGCCCTACCCAGCCTCCAGCTGAGC 1096

RESULT 5  
US-09-535-008-40  
Sequence 40, Application US/09535008  
Patent No. 6465629  
GENERAL INFORMATION:  
APPLICANT: Wong, Alexander K.C.  
APPLICANT: Tavligian, Sean V.  
APPLICANT: Teng, David H.F.  
TITLE OF INVENTION: Bcrl IS A TUMOR SUPPRESSOR THAT IS MUTATED IN PROSTATE  
TITLE OF INVENTION: AND OTHER CANCER TYPES  
FILE REFERENCE: 2318-259  
CURRENT APPLICATION NUMBER: US/09/535.008  
CURRENT FILING DATE: 2000-03-23  
EARLIER APPLICATION NUMBER: U.S. 60/125.806  
EARLIER FILING DATE: 1999-03-23  
NUMBER OF SEQ ID NOS: 77  
SOFTWARE: Patencin Ver. 2.0  
SEQ ID NO 40  
LENGTH: 724  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-535-008-40

Alignment Scores:

Pred. No.: 109 Length: 724  
Score: 46.00 Matches: 8  
Percent Similarity: 84.62% Conservative: 3  
Best Local Similarity: 61.34% Mismatches: 2  
Query Match: 61.33% Indels: 0  
DB: 4 Gaps: 0

US-10-075-846-4\_COPY\_192\_207 (1-16) x US-09-535-008-40 (1-724)  
QY 4 lleucySerrProleupProSerleuSerleuSerVal 16  
DB 681 GTTCTGTCTCCCGCCGCTCCCTCCCTTCACCTACCTC 719

RESULT 6  
US-09-179-221D-3/C  
Sequence 3, Application US/09179221D  
Patent No. 6291168  
GENERAL INFORMATION:  
APPLICANT: Musso, Richard  
TITLE OF INVENTION: NUCLEIC ACID SEQUENCES DIAGNOSTIC FOR PATHOGENIC E.  
TITLE OF INVENTION: COLI 0157, METHODS OF IDENTIFICATION AND KIT THEREFOR  
FILE REFERENCE: 922.6496P  
CURRENT APPLICATION NUMBER: US/09/179.221D  
CURRENT FILING DATE: 1998-10-27  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: No. 6291168epad  
SEQ ID NO 3  
LENGTH: 1331  
TYPE: DNA  
ORGANISM: E. coli  
US-09-179-221D-3

Alignment Scores:  
Pred. No.: 225 Length: 1331  
Score: 46.00 Matches: 8  
Percent Similarity: 80.00% Conservative: 4  
Best Local Similarity: 53.33% Mismatches: 3  
Query Match: 61.33% Indels: 0  
DB: 4 Gaps: 0

US-10-075-846-4\_COPY\_192\_207 (1-16) x US-09-179-221D-3 (1-1331)  
QY 2 SerrlelleucySerrProleupProSerleuSerleuSerVal 16  
DB 102 ACATCAATCATCTCTTCACATGACTAGTCTAAACTAGATT 58

RESULT 7  
US-09-453-702B-70  
Sequence 70, Application US/09453702B  
Patent No. 6365723  
GENERAL INFORMATION:  
APPLICANT: Blattner, Frederick R.  
APPLICANT: Perna, Nicole T.  
APPLICANT: Plunkett, Guy  
TITLE OF INVENTION: No. 6365723el Sequences of E. coli 0157  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Quarles & Brady  
ADDRESS: 1 South Plinckney Street  
CITY: Madison  
STATE: WI  
COUNTRY: US  
ZIP: 53701-2113  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb storage  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 8.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/453.702B  
FILING DATE: 03-Dec-1999

```
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/110,955
FILING DATE: 04-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 960296.95017
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 251-5000
TELEFAX: (608) 251-9166
INFORMATION FOR SEQ ID NO: 70:
SEQUENCE CHARACTERISTICS:
LENGTH: 3191
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 70:
US-09-453-702B-70

Alignment Scores:
Pred. No.: 635 Length: 3191
Score: 46.00 Matches: 8
Percent Similarity: 80.008 Conservative: 4
Best Local Similarity: 53.338 Mismatches: 3
Query Match: 61.338 Indels: 0
DB: 4 Gaps: 0

US-10-075-846-4_COPY_192_207 (1-16) x US-09-453-702B-70 (1-3191)
QY 2 SerSerIleuGysSerProleuProSerIleuSerIleuSerVal 16
Db 547 ACATCAATCATCTGTTCTCAATGACTAGCTAAAGACAGTATT 591
RESULT 8
US-08-254-358-1
Sequence 1, Application US/08254358
Patent No. 5658785
GENERAL INFORMATION:
APPLICANT: Johnson, Phillip R.
TITLE OF INVENTION: Adeno-Associated Virus Materials and
METHODS
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 S. Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/254,358
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 5658785and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 31975
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4680 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
```

```
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-254-358-1

Alignment Scores:
Pred. No.: 1e+03 Length: 4680
Score: 46.00 Matches: 10
Percent Similarity: 73.338 Conservative: 1
Best Local Similarity: 66.678 Mismatches: 4
Query Match: 61.338 Indels: 0
DB: 1 Gaps: 0

US-10-075-846-4_COPY_192_207 (1-16) x US-08-254-358-1 (1-4680)
QY 1 SerSerIleuGysSerProleuProSerIleuSerIleuSer 15
Db 775 ACATCCCAATTAATCTGCTCCCAAAACCCAGCTGACCTCAGT 819
RESULT 9
US-08-475-391-1
Sequence 1, Application US/08475391
Patent No. 5786211
GENERAL INFORMATION:
APPLICANT: Johnson, Phillip R.
TITLE OF INVENTION: Adeno-Associated Virus Materials and
METHODS
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 S. Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,391
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/254,358
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: No. 5786211and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 31975
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4680 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-475-391-1

Alignment Scores:
Pred. No.: 1e+03 Length: 4680
Score: 46.00 Matches: 10
Percent Similarity: 73.338 Conservative: 1
Best Local Similarity: 66.678 Mismatches: 4
Query Match: 61.338 Indels: 0
DB: 1 Gaps: 0

US-10-075-846-4_COPY_192_207 (1-16) x US-08-475-391-1 (1-4680)
```

QY 1 SerSerSerIleuGysSerProleuProSerIleuSerIleuSer 15  
Db 775 ACATCCCAATTACTTGCTCCCAAAACCAACGAGCTGCTCCAGT 819

RESULT 10  
US-08-709-609-1  
Sequence 1, Application US/08709609  
Patent No. 5858775  
GENERAL INFORMATION:  
APPLICANT: Johnson, Phillip R.  
TITLE OF INVENTION: Adeno-Associated Virus Materials and  
METHODS  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 S. Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/709,609  
FILING DATE:  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: No. 5858775and, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 31975  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 474-6300  
TELEFAX: (312) 474-0448  
TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4680 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-709-609-1

Alignment Scores:  
Pred. No.: 1e+03 Length: 4680  
Score: 46.00 Matches: 10  
Percent Similarity: 73.33% Conservative: 1  
Best Local Similarity: 66.67% Mismatches: 4  
Query Match: 61.33% Indels: 0  
DB: 2 Gaps: 0

US-10-075-846-4\_COPY\_192\_207 (1-16) x US-08-709-609-1 (1-4680)

QY 1 SerSerSerIleuGysSerProleuProSerIleuSerIleuSer 15  
Db 775 ACATCCCAATTACTTGCTCCCAAAACCAACGAGCTGCTCCAGT 819

RESULT 11  
PCT-US95-07178-1  
Sequence 1, Application PC/TUS9507178  
GENERAL INFORMATION:  
APPLICANT: Johnson, Phillip R.  
TITLE OF INVENTION: Adeno-Associated Virus Materials and  
METHODS  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 S. Wacker Drive  
CITY: Chicago  
STATE: Illinois

COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/07178  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Noland, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 31975  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 474-6300  
TELEFAX: (312) 474-0448  
TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4680 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
PCT-US95-07178-1

Alignment Scores:  
Pred. No.: 1e+03 Length: 4680  
Score: 46.00 Matches: 10  
Percent Similarity: 73.33% Conservative: 1  
Best Local Similarity: 66.67% Mismatches: 4  
Query Match: 61.33% Indels: 0  
DB: 5 Gaps: 0

US-10-075-846-4\_COPY\_192\_207 (1-16) x PCT-US95-07178-1 (1-4680)

QY 1 SerSerSerIleuGysSerProleuProSerIleuSerIleuSer 15  
Db 775 ACATCCCAATTACTTGCTCCCAAAACCAACGAGCTGCTCCAGT 819

RESULT 12  
US-08-331-384-2/C  
Sequence 2, Application US/08331384  
Patent No. 5856152  
GENERAL INFORMATION:  
APPLICANT: Wilson, James M.  
APPLICANT: Kelley, William M.  
APPLICANT: Fisher, Krishna J.  
TITLE OF INVENTION: Hybrid Adenovirus-AAV Vector and Methods  
OF USE THEREOF  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Howson and Howson  
STREET: Spring House Corporate Cntr, P.O. Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/331,384  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Bak, Mary E.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: UPNG114905A

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9200  
TELEFAX: 215-540-5818  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4910 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: unknown  
MOLECULE TYPE: CDNA  
US-08-331-384-2

Alignment Scores:  
Pred. No.: 1.06e+03 Length: 4910  
Score: 46.00 Matches: 10  
Percent Similarity: 73.33% Conservative: 1  
Best Local Similarity: 66.67% Mismatches: 4  
Query Match: 61.33% Indels: 0  
Gaps: 0  
DB: 2

US-10-075-846-4\_COPY\_192\_207 (1-16) x US-08-331-384-2 (1-4910)

Oy 1 SerSerSerIleLeuCySerProLeuProSerIleuSer 15  
Db 2062 ACATCCCAATTACTGCTCCCAAAACCGCTGAGCTCCAGT 2018

RESULT 13  
US-08-836-087-2/c  
Sequence 2, Application US/08836087  
Patent No. 5871982  
GENERAL INFORMATION:  
APPLICANT: Trustees of University of Pennsylvania  
APPLICANT: Wilson, James M.  
APPLICANT: Kelley, William M.  
APPLICANT: Fisher, Krishna J.  
TITLE OF INVENTION: Hybrid Adenovirus-AAV Vector and  
TITLE OF INVENTION: Methods of Use Thereof  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Howson and Howson  
STREET: Spring House Corporate Cntr, PO Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/836.087  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/331.384  
FILING DATE: 28-OCT-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Bak, Mary E.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: GNVPN.007PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9200  
TELEFAX: 215-540-5818  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4910 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: unknown  
MOLECULE TYPE: CDNA  
US-08-836-087-2

Alignment Scores:  
Pred. No.: 1.06e+03 Length: 4910  
Score: 46.00 Matches: 10  
Percent Similarity: 73.33% Conservative: 1  
Best Local Similarity: 66.67% Mismatches: 4  
Query Match: 61.33% Indels: 0  
Gaps: 0  
DB: 2

US-10-075-846-4\_COPY\_192\_207 (1-16) x US-08-836-087-2 (1-4910)

Oy 1 SerSerSerIleLeuCySerProLeuProSerIleuSer 15  
Db 2062 ACATCCCAATTACTGCTCCCAAAACCGCTGAGCTCCAGT 2018

RESULT 14  
US-09-246-320-2/c  
Sequence 2, Application US/09246320  
Patent No. 6251677  
GENERAL INFORMATION:  
APPLICANT: Trustees of University of Pennsylvania  
APPLICANT: Wilson, James M.  
APPLICANT: Kelley, William M.  
APPLICANT: Fisher, Krishna J.  
TITLE OF INVENTION: Hybrid Adenovirus-AAV Vector and  
TITLE OF INVENTION: Methods of Use Thereof  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Howson and Howson  
STREET: Spring House Corporate Cntr, PO Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/246.320  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/836.087  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Bak, Mary E.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: GNVPN.007PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9200  
TELEFAX: 215-540-5818  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4910 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: unknown  
MOLECULE TYPE: CDNA  
US-09-246-320-2

Alignment Scores:  
Pred. No.: 1.06e+03 Length: 4910  
Score: 46.00 Matches: 10  
Percent Similarity: 73.33% Conservative: 1  
Best Local Similarity: 66.67% Mismatches: 4  
Query Match: 61.33% Indels: 0  
Gaps: 0  
DB: 4

US-10-075-846-4\_COPY\_192\_207 (1-16) x US-09-246-320-2 (1-4910)

Oy 1 SerSerSerIleLeuCySerProLeuProSerIleuSer 15  
Db 2062 ACATCCCAATTACTGCTCCCAAAACCGCTGAGCTCCAGT 2018

Db 2062 ACATCCCAATTACTTGCTCCCAAAACCCAGCCTGAGCTCCAGT 2018

RESULT 15  
US-09-546-738-2/C  
Sequence 2, Application US/09546738  
Patent No. 6387368  
GENERAL INFORMATION:  
APPLICANT: Trustees of University of Pennsylvania  
Wilson, James M.  
Kelley, William M.  
Fisher, Krishna J.  
TITLE OF INVENTION: Hybrid Adenovirus-AAV Vector and  
Methods of Use Thereof  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Howson and Howson  
STREET: Spring House Corporate Cntr, PO Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/546,738  
FILING DATE: 11-Apr-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/246,320  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Bak, Mary E.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: GNVPN.007PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9200  
TELEFAX: 215-540-5818  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4910 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: unknown  
MOLECULE TYPE: cDNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-546-738-2

Alignment Scores:  
Pred. No.: 1.06e+03 Length: 4910  
Score: 46.00 Matches: 10  
Percent Similarity: 73.33% Conservative: 1  
Best Local Similarity: 66.67% Mismatches: 4  
Query Match: 61.33% Indels: 0  
DB: 4 Gaps: 0

US-10-075-846-4\_COPY\_192\_207 (1-16) x US-09-546-738-2 (1-4910)

Qy 1 SerSerSerIleLeuCySerProLeuProSerLeuSerLeuSer 15  
:::||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 2062 ACATCCCAATTACTTGCTCCCAAAACCCAGCCTGAGCTCCAGT 2018

Search completed: July 1, 2003, 00:57:37  
Job time : 12.4004 secs



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 30, 2003, 23:07:05 ; Search time 8.51902 Seconds  
(without alignments)  
2788.018 Million cell updates/sec

Title: US-10-075-846-4\_COPY\_192\_207

Perfect score: 75

Sequence: 1 SSSILCSPLPSLSLV 16

Scoring table: BIOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 1055720 segs, 742224136 residues

Total number of hits satisfying chosen parameters: 2111440

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n model -DEV=xlp  
-O=/cgn2\_1/USPRO/spool/US10075846/runat\_25062003\_163649\_5181/app.query.fasta.1.782  
-DB=Published,Applications\_NA -OFMT=fastap -SUFFIX=p2n.rmpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blsum62  
-TRANS=Numan40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pcr -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=plc -NORM=ext -HEADSIZE=500 -MINLEN=0  
-MALEN=200000000 -USER=US10075846 @CGN.1.1.102 @runat\_25062003\_163649\_5181  
-NCPU=6 -ICPU=3 -NO\_MAP -LARGQUERY -NEG\_SCORES=0 -WAIT -DSBLOCK=100  
-LONGLOC -DEV=TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Published Applications\_NA:\*

1: /cgn2\_6/prodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/prodata/2/pubpna/PCR\_NEM\_PUB.seq:\*  
3: /cgn2\_6/prodata/2/pubpna/US06\_NEM\_PUB.seq:\*  
4: /cgn2\_6/prodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/prodata/2/pubpna/US07\_NEM\_PUB.seq:\*  
6: /cgn2\_6/prodata/2/pubpna/PCRUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/prodata/2/pubpna/US08\_NEM\_PUB.seq:\*  
8: /cgn2\_6/prodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/prodata/2/pubpna/US09\_NEM\_PUB.seq:\*  
10: /cgn2\_6/prodata/2/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/prodata/2/pubpna/US10\_NEM\_PUB.seq:\*  
12: /cgn2\_6/prodata/2/pubpna/US10\_PUBCOMB.seq:\*  
13: /cgn2\_6/prodata/2/pubpna/US60\_NEM\_PUB.seq:\*  
14: /cgn2\_6/prodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	75	100.0	1640 9	US-10-075-846-3
2	53	70.7	2982 9	US-08-938-842A-93
3	53	70.7	3301 10	US-09-954-456-554
4	52	69.3	90 10	US-09-864-761-30015

C 5	52	69.3	597 10	US-09-864-761-13469	Sequence 13469, A
C 6	49	65.3	249487 9	US-10-026-188-3	Sequence 3, Appl1
C 7	48	64.0	497 10	US-09-783-590-11273	Sequence 11273, A
C 8	48	64.0	576 10	US-09-974-300-865	Sequence 865, App
C 9	48	64.0	720 9	US-10-138-846-12894	Sequence 12894, A
C 10	48	64.0	791 9	US-10-198-846-7241	Sequence 7241, App
C 11	48	64.0	5399 9	US-10-225-567A-343	Sequence 343, App
C 12	48	64.0	10468 9	US-09-764-891-8976	Sequence 8976, App
C 13	48	64.0	10468 9	US-09-764-891-8976	Sequence 8976, App
C 14	47	62.7	263 9	US-09-918-985-8979	Sequence 8979, App
C 15	47	62.7	381 10	US-09-983-965-3940	Sequence 3940, App
C 16	47	62.7	771 9	US-09-764-891-725	Sequence 725, App
C 17	47	62.7	4000 10	US-09-880-107-3430	Sequence 3430, App
C 18	47	62.7	110079 9	US-10-175-523-96	Sequence 96, Appl
C 19	47	62.7	174424 10	US-09-967-768A-314	Sequence 314, App
C 20	47	62.7	465237 10	US-09-933-267A-1	Sequence 1, Appl1
C 21	46	61.3	231 10	US-09-864-761-17984	Sequence 17984, A
C 22	46	61.3	254 9	US-09-728-444-304	Sequence 304, App
C 23	46	61.3	289 9	US-09-991-936-556	Sequence 556, App
C 24	46	61.3	408 9	US-09-918-995-23458	Sequence 23458, A
C 25	46	61.3	424 10	US-09-783-590-2728	Sequence 2728, App
C 26	46	61.3	479 10	US-09-864-761-1221	Sequence 1221, App
C 27	46	61.3	1363 10	US-09-880-192-27	Sequence 27, Appl
C 28	46	61.3	1591 10	US-09-925-300-290	Sequence 290, App
C 29	46	61.3	1611 9	US-09-792-630-16	Sequence 16, Appl
C 30	46	61.3	1611 9	US-09-792-630-18	Sequence 18, Appl
C 31	46	61.3	1611 9	US-10-080-376-16	Sequence 16, Appl
C 32	46	61.3	1611 9	US-10-080-376-18	Sequence 18, Appl
C 33	46	61.3	1611 9	US-09-953-351-16	Sequence 16, Appl
C 34	46	61.3	1611 9	US-09-953-351-18	Sequence 18, Appl
C 35	46	61.3	1611 9	US-10-082-671-22	Sequence 22, Appl
C 36	46	61.3	1611 9	US-10-082-671-24	Sequence 24, Appl
C 37	46	61.3	1611 9	US-10-097-100-16	Sequence 16, Appl
C 38	46	61.3	1611 9	US-10-097-100-18	Sequence 18, Appl
C 39	46	61.3	1866 9	US-09-792-630-12	Sequence 12, Appl
C 40	46	61.3	1866 9	US-09-792-630-4	Sequence 4, Appl1
C 41	46	61.3	1866 9	US-10-080-376-2	Sequence 2, Appl1
C 42	46	61.3	1866 9	US-10-080-376-4	Sequence 4, Appl1
C 43	46	61.3	1866 9	US-09-953-351-2	Sequence 2, Appl1
C 44	46	61.3	1866 9	US-09-953-351-4	Sequence 4, Appl1
C 45	46	61.3	1866 9	US-10-082-671-8	Sequence 8, Appl1

## ALIGNMENTS

RESULT 1  
US-10-075-846-3  
: Sequence 3, Application US/10075846  
: Publication No. US20030032608A1  
: GENERAL INFORMATION:  
: APPLICANT: Bristol-Myers Squibb Company  
: TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING A NOVEL GLYCINE RECEPTOR ALPHA SUBUNIT  
: FILE REFERENCE: IN THE GASTROINTESTINAL TRACT, HGRA4, and SPLICE VARIANT THERE  
: CURRENT APPLICATION NUMBER: US/10/075, 846  
: PRIOR FILING DATE: 2002-02-13  
: PRIOR APPLICATION NUMBER: US 60/269, 535  
: NUMBER OF SEQ ID NOS: 81  
: SOFTWARE: PatentIn version 3.0  
: SEQ ID NO 3  
: LENGTH: 1640  
: TYPE: DNA  
: ORGANISM: homo sapiens  
: FEATURE:  
: NAME/KEY: CDS  
: LOCATION: (1)..(1293)  
US-10-075-846-3

Alignment Scores:  
Pred. No.: 0.00593  
Score: 75.00  
Percent Similarity: 100.00%

Length: 1640  
Matches: 16  
Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-10-075-846-4\_COPY\_192\_207 (1-16) x US-10-075-846-3 (1-1640)

OY 1 SerSerSerIleLeuGysSerProLeuProSerIleuSerIleuSerVal 16  
DB 574 AGCTCATCATCATCTGTGACGCCCTCTGCATCTCTGTCATCTTCAGTT 621

RESULT 2  
US-09-938-842A-93/C  
Sequence 93, Application US/09938842A  
Patent No. US20020160378A1  
GENERAL INFORMATION:  
APPLICANT: Harper, Jeff  
APPLICANT: Kreps, Joel  
APPLICANT: Wang, Xun  
APPLICANT: Zhu, Tong  
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
FILE REFERENCE: SCRIPI300-3  
CURRENT APPLICATION NUMBER: US/09/938,842A  
CURRENT FILING DATE: 2001-08-24  
PRIOR APPLICATION NUMBER: US 60/227,866  
PRIOR FILING DATE: 2000-08-24  
PRIOR APPLICATION NUMBER: US 60/264,647  
PRIOR FILING DATE: 2001-01-16  
PRIOR APPLICATION NUMBER: US 60/300,111  
PRIOR FILING DATE: 2001-06-22  
NUMBER OF SEQ ID NOS: 5379  
SEQ ID NO 93  
LENGTH: 2982  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
US-09-938-842A-93

Alignment Scores:  
Pred. No.: 40.1 Length: 2982  
Score: 53.00 Matches: 10  
Percent Similarity: 81.25% Conservative: 3  
Best Local Similarity: 62.50% Mismatches: 3  
Query Match: 70.67% Indels: 0  
Gaps: 0

US-10-075-846-4\_COPY\_192\_207 (1-16) x US-09-938-842A-93 (1-2982)

OY 1 SerSerSerIleLeuGysSerProLeuProSerIleuSerIleuSerVal 16  
DB 2893 TCTTATTGAGTAGCTGTGCACCTCTCCCTCAATCTCTGTATCAGTT 2846

RESULT 3  
US-09-954-456-554  
Sequence 554, Application US/09954456  
Patent No. US20020115057A1  
GENERAL INFORMATION:  
APPLICANT: Young, Paul  
TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand  
FILE REFERENCE: 689290-76  
CURRENT APPLICATION NUMBER: US/09/954,456  
CURRENT FILING DATE: 2001-09-18  
PRIOR APPLICATION NUMBER: US/60/233,617  
PRIOR FILING DATE: 2000-09-18  
PRIOR APPLICATION NUMBER: US/60/234,052  
PRIOR FILING DATE: 2000-09-20  
PRIOR APPLICATION NUMBER: US/60/234,923  
PRIOR FILING DATE: 2000-09-25  
PRIOR APPLICATION NUMBER: US/60/235,134  
PRIOR FILING DATE: 2000-09-25  
PRIOR APPLICATION NUMBER: US/60/235,637  
PRIOR FILING DATE: 2000-09-26  
PRIOR APPLICATION NUMBER: US/60/235,638

PRIOR FILING DATE: 2000-09-26  
PRIOR APPLICATION NUMBER: US/60/235,711  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: US/60/235,720  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: US/60/235,840  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: US/60/235,863  
PRIOR FILING DATE: 2000-09-27  
NUMBER OF SEQ ID NOS: 2276  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 554  
LENGTH: 3301  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc-feature  
OTHER INFORMATION: n-a,t,g or c  
US-09-954-456-554

Alignment Scores:  
Pred. No.: 44.8 Length: 3301  
Score: 53.00 Matches: 12  
Percent Similarity: 92.86% Conservative: 1  
Best Local Similarity: 85.71% Mismatches: 1  
Query Match: 70.67% Indels: 0  
Gaps: 0

US-10-075-846-4\_COPY\_192\_207 (1-16) x US-09-954-456-554 (1-3301)

OY 2 SerSerSerIleLeuGysSerProLeuProSerIleuSerIleuSer 15  
DB 1837 TCTTCCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 1878

RESULT 4  
US-09-864-761-30015/C  
Sequence 30015, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharron G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO  
FILE REFERENCE: Aegmiga-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263,6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662

```

: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
: SEQ ID NO 30015
: LENGTH: 90
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AL138761.3
: OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.7
: OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.6
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
: OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.7
: OTHER INFORMATION: NT HIT: AB021490.2, EVALUATE 7.40e-01
: OTHER INFORMATION: EST_HUMAN HIT: BE612820.1, EVALUATE 1.10e+00
US-09-864-761-30015
```

```

Alignment Scores:
Pred. No.: 1.25 Length: 90
Score: 52.00 Matches: 10
Percent Similarity: 92.31% Conservative: 2
Best Local Similarity: 76.92% Mismatches: 1
Query Match: 69.33% Indels: 0
DB: 10 Gaps: 0
```

US-10-075-846-4\_COPY\_192\_207 (1-16) x US-09-864-761-30015 (1-90)

```
OY 3 SerileuCySerProleupProserleuSerleuSer 15
||||| ||||||||| ||||||||| |||||||||
Db 81 TCGCATCTTGGCTGCTCTCTCCCTCTCTCTCTCTCC 43
```

```

RESULT 5
US-09-864-761-13469/c
: Sequence 13469, Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
: FILE REFERENCE: Aecmics-X-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263.6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
```

```

: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
: SEQ ID NO 13469
: LENGTH: 597
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AL138761.3
: OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.7
: OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.6
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
: OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.7
US-09-864-761-13469
```

```

Alignment Scores:
Pred. No.: 9.97 Length: 597
Score: 52.00 Matches: 10
Percent Similarity: 92.31% Conservative: 2
Best Local Similarity: 76.92% Mismatches: 1
Query Match: 69.33% Indels: 0
DB: 10 Gaps: 0
```

US-10-075-846-4\_COPY\_192\_207 (1-16) x US-09-864-761-13469 (1-597)

```
OY 3 SerileuCySerProleupProserleuSerleuSer 15
||||| ||||||||| ||||||||| |||||||||
Db 443 TCGCATCTTGGCTGCTCTCTCTCCCTCTCTCTCTCTCC 405
```

```

RESULT 6
US-10-026-188-3
: Sequence 3, Application US/10026188
: Patent No. US20020164645A1
: GENERAL INFORMATION:
: APPLICANT: Zuker, Charles S.
: APPLICANT: Zhang, Yifeng
: APPLICANT: The Regents of the University of California
: TITLE OF INVENTION: Assays for Taste Receptor Cell Specific
: FILE REFERENCE: 02307E-114910US
: CURRENT APPLICATION NUMBER: US/10/026,188
: CURRENT FILING DATE: 2001-12-21
: PRIOR APPLICATION NUMBER: US 60/259,379
: PRIOR FILING DATE: 2000-12-29
: NUMBER OF SEQ ID NOS: 8
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 3
: LENGTH: 249487
: TYPE: DNA
: ORGANISM: Mus musculus
: FEATURE:
: OTHER INFORMATION: mouse genomic region containing Itprcs
US-10-026-188-3

Alignment Scores:
Pred. No.: 2.25e+04 Length: 249487
```



```
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974.300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680.598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279.526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 865
; LENGTH: 576
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-974-300-865

Alignment Scores:
Pred. No.: 42.3      Length: 576
Score: 48.00      Matches: 9
Percent Similarity: 85.71%      Conservative: 3
Best Local Similarity: 64.29%      Mismatches: 2
Query Match: 64.00%      Indels: 0
DB: 10      Gaps: 0

US-10-075-846-4_COPY_192_207 (1-16) x US-09-974-300-865 (1-576)
Oy 2 SerSerIleuCySerProleuProserIleuSerIleuSer 15
DB 557 GCTTCTACTATATCTCTACCGCTCTCTCTTTCAATCGCG 516

RESULT 9
US-10-198-846-12894/C
; Sequence 12894, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198.846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306.220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 12894
; LENGTH: 720
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc-feature
; LOCATION: 1, 2, 3, 716, 717, 718, 719, 720
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-12894

Alignment Scores:
Pred. No.: 54      Length: 720
Score: 48.00      Matches: 8
Percent Similarity: 80.00%      Conservative: 4
Best Local Similarity: 53.33%      Mismatches: 3
Query Match: 64.00%      Indels: 0
DB: 9      Gaps: 0

US-10-075-846-4_COPY_192_207 (1-16) x US-10-198-846-12894 (1-720)
Oy 1 SerSerIleuCySerProleuProserIleuSerIleuSer 15
DB 316 ACTACGGCTCATCTCTCGCGGCCCATCAGCCTCATGCGC 272

RESULT 10
```

```
US-10-198-846-7241/C
; Sequence 7241, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198.846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306.220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 7241
; LENGTH: 791
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc-feature
; LOCATION: 282, 300, 554, 592, 650, 707, 720, 728, 755, 767, 775, 791
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-7241

Alignment Scores:
Pred. No.: 59.9      Length: 791
Score: 48.00      Matches: 8
Percent Similarity: 80.00%      Conservative: 4
Best Local Similarity: 53.33%      Mismatches: 3
Query Match: 64.00%      Indels: 0
DB: 9      Gaps: 0

US-10-075-846-4_COPY_192_207 (1-16) x US-10-198-846-7241 (1-791)
Oy 1 SerSerIleuCySerProleuProserIleuSerIleuSer 15
DB 81 ACTACGGCTCATCTCTCGCGGCCCATCAGCCTCATGCGC 37

RESULT 11
US-10-225-567A-343
; Sequence 343, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: Lifespan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burner, Glenna C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPT
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225.567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257.144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 343
; LENGTH: 5399
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-225-567A-343

Alignment Scores:
Pred. No.: 491      Length: 5399
Score: 48.00      Matches: 9
Percent Similarity: 90.91%      Conservative: 1
Best Local Similarity: 81.82%      Mismatches: 1
Query Match: 64.00%      Indels: 0
DB: 9      Gaps: 0
```

US-10-075-846-4\_COPY\_192\_207 (1-16) x US-10-225-567A-343 (1-5399)

OY 5 SerSettleucysSerProleuproSerleuSer 15

DB 2876 CTGTGCGCCCTTACCCAGCCTCCAGCTGAGC 2908

RESULT 12

US-09-764-891-8976  
; Sequence 8976, Application US/09764891  
; Publication No. US20030077808A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC006  
; CURRENT APPLICATION NUMBER: US/09/764,891  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 10231  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 8976  
; LENGTH: 10468  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-764-891-8976

Alignment Scores:

Pred. No.: 1.01e+03 Length: 10468  
Score: 48.00 Matches: 9  
Percent Similarity: 90.91% Conservative: 1  
Best Local Similarity: 81.82% Mismatches: 1  
Query Match: 64.00% Indels: 0  
DB: 9 Gaps: 0

US-10-075-846-4\_COPY\_192\_207 (1-16) x US-09-764-891-8976 (1-10468)

OY 2 SerSettleucysSerProleuproSerleu 12

DB 4445 AGCTCTGCTCTTGTAGCCCTTACCACTGCTC 4477

RESULT 13

US-09-764-891-8979  
; Sequence 8979, Application US/09764891  
; Publication No. US20030077808A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC006  
; CURRENT APPLICATION NUMBER: US/09/764,891  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 10231  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 8979  
; LENGTH: 10468  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-764-891-8979

Alignment Scores:

Pred. No.: 1.01e+03 Length: 10468  
Score: 48.00 Matches: 9  
Percent Similarity: 90.91% Conservative: 1  
Best Local Similarity: 81.82% Mismatches: 1  
Query Match: 64.00% Indels: 0  
DB: 9 Gaps: 0

US-10-075-846-4\_COPY\_192\_207 (1-16) x US-09-764-891-8979 (1-10468)

OY 2 SerSettleucysSerProleuproSerleu 12

DB 4445 AGCTCTGCTCTTGTAGCCCTTACCACTGCTC 4477

RESULT 14

US-09-918-995-30770

; Sequence 30770, Application US/09918995

; Publication No. US20030073623A1

; GENERAL INFORMATION:

; APPLICANT: HySeq, Inc.  
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
; FROM VARIOUS CDNA LIBRARIES  
; FILE REFERENCE: 20411-756  
; CURRENT APPLICATION NUMBER: US/09/918,995  
; CURRENT FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: US/09/235,076  
; PRIOR FILING DATE: 1999-01-20  
; NUMBER OF SEQ ID NOS: 38054  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 30770  
; LENGTH: 263  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(263)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-918-995-30770

Alignment Scores:

Pred. No.: .26 Length: 263  
Score: 47.00 Matches: 9  
Percent Similarity: 84.62% Conservative: 2  
Best Local Similarity: 69.23% Mismatches: 2  
Query Match: 62.67% Indels: 0  
DB: 9 Gaps: 0

US-10-075-846-4\_COPY\_192\_207 (1-16) x US-09-918-995-30770 (1-263)

OY 3 SerSettleucysSerProleuproSerleuSer 15

DB 68 GCTCTCTGCTCTTCTGCCCCCTTACCACTGCTCTTCT 106

RESULT 15

US-09-983-965-3940/C  
; Sequence 3940, Application US/09983965  
; Patent No. US20020137160A1  
; GENERAL INFORMATION:  
; APPLICANT: Warren, Wesley C.  
; APPLICANT: Tao, Nengbing  
; APPLICANT: Byatt, John C.  
; APPLICANT: Mathialagan, Nagappan  
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND  
; FILE REFERENCE: 37-21(10297)C  
; CURRENT APPLICATION NUMBER: US/09/983,965  
; CURRENT FILING DATE: 2001-10-26  
; PRIOR APPLICATION NUMBER: US 09/465,231  
; PRIOR FILING DATE: 1999-12-15  
; PRIOR APPLICATION NUMBER: US 60/113,678  
; PRIOR FILING DATE: 1998-12-17  
; NUMBER OF SEQ ID NOS: 5912  
; SEQ ID NO 3940  
; LENGTH: 381  
; TYPE: DNA  
; ORGANISM: Bos taurus  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 51-LIB3058-053-Q1-K1-E12  
US-09-983-965-3940

Alignment Scores:

Pred. No.: 39 Length: 381  
Score: 47.00 Matches: 9  
Percent Similarity: 78.57% Conservative: 2  
Best Local Similarity: 64.29% Mismatches: 3  
Query Match: 62.67% Indels: 0  
DB: 10 Gaps: 0



US-10-075-846-4\_COPY\_192\_207 (1-16) x US-09-983-965-3940 (1-381)

QY 2 SerSerIleLeuCySerProLeuProSerLeuSerLeuSer 15

Db 144 AGCGGCATCTTGCTTACCGCTGCCAGCCTTGGAATCAGC 103

Search completed: July 1, 2003, 01:01:55  
Job time : 29.519 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: June 30, 2003, 23:00:29 ; Search time 84.1163 Seconds

(without alignments)  
3080.588 Million cell updates/sec

Title: US-10-075-846-4\_COPY\_192\_207  
Perfect score: 75  
Sequence: 1 SSSILCSPSLPSLSLV 16

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08  
Maximum Match 1008

Listing first 45 summaries

Command line parameters:

-MODL=frame\_plus.p2n.model -DEV=xlp  
-O/cgn2.1/USPTO.spool/US10075846/runat\_25062003.163648.5157/app.query.fasta\_1.782  
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45  
-DOCLALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pico -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10075846.ecgn.1.1.3724.ernat.25062003.163648.5157 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEOBJECT -NEG.SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6  
-Fgapext=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEX=7

Database :

1: em\_estba:\*  
2: em\_estchn:\*  
3: em\_estlin:\*  
4: em\_estlm:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hcc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hcc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vit:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56	74.7	795	17 AG076357	AG076357 Pan trogl
2	55	73.3	310	14 B0484167	B0484167 pmp1c.pko
3	55	73.3	884	12 BE739001	BE739001 601556067
4	55	73.3	1081	12 BG622007	BG622007 602646601
5	54	72.0	128	17 A2485951	A2485951 100313112
6	53	70.7	263	10 BE526547	BE526547 M55F05STM
7	53	70.7	306	9 AA597350	AA597350 29621.fam
8	53	70.7	404	10 AV817296	AV817296 AV817296
9	53	70.7	470	10 BE591974	BE591974 WHE1651-1
10	53	70.7	547	10 AV548451	AV548451 AV548451
11	53	70.7	548	10 AV548438	AV548438 AV548438
12	53	70.7	556	10 AV547337	AV547337 AV547337
13	53	70.7	597	17 A0415379	A0415379 RPT-11-1
14	53	70.7	600	14 B0246138	B0246138 TAB15015F
15	53	70.7	615	13 B0268516	B0268516 B0268516
16	53	70.7	668	13 B0237354	B0237354 B0237354
17	53	70.7	787	17 A2198854	A2198854 SP_1038_B
18	53	70.7	1329	12 BG167249	BG167249 603242511
19	52	69.3	287	13 BM346474	BM346474 Kx71e05.y
20	52	69.3	364	9 A1164587	A1164587 A065P41U
21	52	69.3	443	17 A2713283	A2713283 RPT-24-6
22	52	69.3	445	13 B1449589	B1449589 dae73b10.
23	52	69.3	471	10 AW591971	AW591971 hf36d10.x
24	52	69.3	484	17 A2826655	A2826655 2M0102K22
25	52	69.3	510	12 BF430050	BF430050 256000.MA
26	52	69.3	537	17 A2966794	A2966794 2M0237P22
27	52	69.3	572	17 A2506723	A2506723 1M0348F03
28	52	69.3	656	17 AG041152	AG041152 Pan trogl
29	52	69.3	782	17 BH501609	BH501609 BOHJ27R
30	52	69.3	834	13 B1091085	B1091085 602855289
31	51.5	68.7	654	12 BF691466	BF691466 602247635
32	51	68.0	231	9 AA967038	AA967038 ua39e01.r
33	51	68.0	242	12 BF552695	BF552695 UI-R-CO-1
34	51	68.0	303	10 AW346805	AW346805 29517.MAR
35	51	68.0	344	10 AW344669	AW344669 24424.MAR
36	51	68.0	359	10 BE117765	BE117765 UI-R-B51-
37	51	68.0	422	13 BM388008	BM388008 UI-R-CM1-
38	51	68.0	479	12 BG380127	BG380127 UI-R-CS0-
39	51	68.0	530	9 A1406681	A1406681 EST234968
40	51	68.0	540	13 B1896388	B1896388 fc80d10.x
41	51	68.0	549	10 BE117524	BE117524 UI-R-B51-
42	51	68.0	558	14 B0780785	B0780785 UI-R-FPO-
43	51	68.0	615	14 B0193942	B0193942 UI-R-CM1-
44	51	68.0	643	9 A1176553	A1176553 EST220140
45	51	68.0	650	17 AG053401	AG053401 Pan trogl

#### ALIGNMENTS

RESULT 1  
AG076357/c 795 bp DNA linear GSS 03-NOV-2001  
LOCUS Pan troglodytes DNA, clone: PTB-070115.F, genomic survey sequence.  
ACCESSION AG076357  
VERSION AG076357.1 GI:16628159  
KEYWORDS  
SOURCE  
ORGANISM  
Pan troglodytes male lymphoblast DNA, clone: PTB Chimpanzee Male  
BAC library clone: PTB-070115.F.  
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.  
REFERENCE  
1 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,  
Totoki, Y., Watanabe, H. and Sakaki, Y.

TITLE BAC end sequences of library PTB  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 795)  
AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T.,  
Totoki, Y., Matanabe, H. and Sakaki, Y.  
TITLE Direct Submission  
JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical  
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);  
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
(E-mail: chimpesgsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/  
Tel: 81-45-503-9111, Fax: 81-45-503-9170)  
Clones are derived from the chimpanzee BAC library PTB. This BAC end  
was generated during the Rad process and may have higher chance of  
clone tracking errors.  
PRIMERS  
Sequencing: -21M13  
LIBRARY  
Vector : pKS145  
R Site 1 : SacI  
R Site 2 : SacI  
Location/Qualifiers  
1. 795  
/organism="Pan troglodytes"  
/db\_xref="taxon:9598"  
/clone="PTB-070115.F"  
/sex="male"  
/cell\_type="lymphoblast"  
/clone\_lib="PTB chimpanzee Male BAC Library"  
BASE COUNT 201 a 115 c 325 g 149 t 5 others  
ORIGIN  
Alignment Scores:  
Pred. No.: 765 Length: 795  
Score: 56.00 Matches: 11  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 91.67% Mismatches: 0  
Query Match: 74.67% Indels: 0  
Gaps: 0  
DB: 17  
US-10-075-846-4\_COPY\_192\_207 (1-16) x AG076357 (1-795)  
QY 3 SerileuCySerProleupProSerleuSerleu 14  
|||||  
Db 712 TCTATATTTATGTTCCCTCCCTCTCTCACGCTA 677  
RESULT 2 310 bp mRNA linear EST 07-JUN-2002  
LOCUS BQ484167  
DEFINITION pmplc.pk002.b9 Chicken Macrophage cDNA library Gallus gallus cDNA  
ACCESSION BQ484167  
VERSION BQ484167.1 GI:21328786  
KEYWORDS EST.  
SOURCE chicken.  
ORGANISM Gallus gallus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
Phasianinae; Gallus.  
REFERENCE 1 (bases 1 to 310)  
AUTHORS Keeler, C.  
TITLE ESTs from Chicken Macrophages  
JOURNAL Unpublished (2002)  
CONTACT: Calvin Keeler  
University of Delaware  
Townsend Hall, Newark, DE 19717, USA  
Tel: 302-831-6473  
Fax: 302-831-2822  
Email: ckeeler@udel.edu  
FEATURES  
source  
1. 310  
/organism="Gallus gallus"  
/db\_xref="taxon:9031"  
/clone="pmplc.pk002.b9"

/clone\_lib="Chicken Macrophage cDNA library"  
/sex="Male and Female"  
/tissue\_type="LPS-stimulated HD-11 macrophage cells"  
/lab\_host="E.coli XL-10 Blue"  
/note="Vector: pBluescript II-SK"  
BASE COUNT 74 a 119 c 56 g 60 t 1 others  
ORIGIN  
Alignment Scores:  
Pred. No.: 278 Length: 310  
Score: 55.00 Matches: 11  
Percent Similarity: 81.25% Conservative: 2  
Best Local Similarity: 68.75% Mismatches: 3  
Query Match: 73.33% Indels: 0  
Gaps: 0  
DB: 14  
US-10-075-846-4\_COPY\_192\_207 (1-16) x BQ484167 (1-310)  
QY 1 SerSerleuCySerProleupProSerleuSerleuSerVal 16  
|||||  
Db 196 TCTCTCTCAACCTCTGCTCCCTCCCAACCTCCGCTCCATC 243  
RESULT 3  
LOCUS BE739001/c 884 bp mRNA linear EST 15-SEP-2000  
DEFINITION 601556067F1 NIH\_MGC\_58 Homo sapiens cDNA clone IMAGE:3825809 5',  
mRNA sequence.  
ACCESSION BE739001  
VERSION BE739001.1 GI:10152993  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 884)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
CONTACT: Robert Strausberg, Ph.D.  
Email: cga@bbs.fda.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: CLONTECH Laboratories, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/ULNL at:  
http://image.llnl.gov  
Plate: L10M492 row: f column: 18  
High quality sequence stop: 555.  
FEATURES  
source  
1. 884  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3825809"  
/clone\_lib="NIH\_MGC\_58"  
/tissue\_type="hypertrophoma"  
/lab\_host="DH10B (T1 phage-resistant)"  
/note="Organ: kidney; Vector: pDNR-LIB (Clontech); Site\_1:  
SfiI (ggcgccatggcc); Site\_2: SfiI (ggcgccatggcc);  
Double-stranded cDNA was prepared from cell line RNA. 5'  
and 3' adaptors were used in cloning as follows: 5'  
adaptor sequence: 5'-CAGGCGCATTTATGCGC-3' and 3' adaptor  
sequence: 5'-ATTCTAGACGCGCGCGCATG-dT(30)BN-3'  
(where B = A, C, G, or T). Average  
insert size 1.35 kb (range 0.9-4.0 kb). 15/15 colonies  
contained inserts by PCR. This library was enriched for  
full-length clones and was constructed by Clontech  
Laboratories (Palo Alto, CA)."  
BASE COUNT 223 a 208 c 280 g 173 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.22e+03 Length: 884



```

US-10-075-846-4_COPY_192_207 (1-16) x AA597350 (1-128)
QY
2 SerSettleucySerProleuProSerleuSerleuServal 16
||||| ||| |||||||||||||||||||||:
Db
2 TCCTCTTCTACTCTCTCCCTCCCTCTTCTTCTCCCTG 46

RESULT 6
BE526547 263 bp mRNA linear EST 19-MAR-2001
LOCUS BE526547/c
DEFINITION M55F0557M Arabidopsis developing seed Arabidopsis thaliana cDNA
ACCESSION BE526547
VERSION BE526547.1 GI:9784525
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 263)
White,J.A., Todd,J., Newman,T., Focks,N., Girke,T., Martinez de
Ilarduya,O., Jaworski,J.G., Ohlrogge,J. and Banning,C.
A new set of Arabidopsis expressed sequence tags from developing
seeds. The metabolic pathway from carbohydrates to seed oil
Plant Physiol. 124 (4), 1382-1394 (2000)
20567808
JOURNAL Contact: Banning, C
MEDLINE Dept. of Biochemistry & Molecular Biology
COMMENT Michigan State University
224 Biochemistry, Michigan State University, East Lansing, MI 48824
Tel: 517 355 1609
Fax: 517 353 9334
Email: banning@msu.edu
Clones were originally prepared at Michigan State University.
Arabidopsis Biologically Resource Center, The Ohio State University,
309 Botany & Zoology Bldg., 1735 Neil Avenue, Columbus, OH 43210
USA, FAX: 6142920603 TEL: 6142929371.
Location/Qualifiers
FEATURES
source
1.263
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="600035219R1"
/clone_lib="Arabidopsis developing seed"
/tissue_type="seed"
/dev_stage="5-13 days after flowering"
/lab_host="E.coli"
/note="Organ: Developing seed; Vector: pBluescript SK-;
Site_1: EcoRI; Site_2: XhoI"
Site_1: 38 c
Site_2: 61 t
BASE COUNT 85 a 38 c 79 g . 61 t
ORIGIN

Alignment Scores:
Pred. No.: 416 Length: 263
Score: 53.00 Matches: 10
Percent Similarity: 81.258 Conservative: 3
Best Local Similarity: 62.508 Mismatches: 3
Query Match: 70.678 Indels: 0
DB: AV817296 Gaps: 0

US-10-075-846-4_COPY_192_207 (1-16) x BE526547 (1-263)
QY
1 SerSettleucySerProleuProSerleuSerleuServal 16
||||| ||| |||||||||||||||||||||:
Db
130 TCTATTACAGTAGCTTCTCCACCTCTCCCTCAATCTCTGATCACTT 83

RESULT 7
AA597350 306 bp mRNA linear EST 19-SEP-1997
LOCUS AA597350/c
DEFINITION 29621 Lambda-PRL2 Arabidopsis thaliana cDNA clone 196B17M4, mRNA
sequence.
ACCESSION AA597350
AUTHORS
```

```

VERSION AA597350.1 GI:2412773
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 306)
Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H., McIntosh
,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomashow,M., Retzel
,E. and Somerville,C.
Genes galore: a summary of methods for accessing results from
large-scale partial sequencing of anonymous Arabidopsis cDNA clones
Plant Physiol. 106, 1241-1255 (1994)
95148729
JOURNAL Contact: Thomas Newman
MEDLINE MSU-DOE Plant Research Laboratory
COMMENT Michigan State University
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22333ct@elm.cl.msu.edu
Seq primer: 5' end dye terminator primer closer than T7.
Location/Qualifiers
FEATURES
source
1.306
/organism="Arabidopsis thaliana"
/strain="var columbia"
/db_xref="taxon:3702"
/clone="196B17M4"
/clone_lib="Lambda-PRL2"
/note="Vector: lambda Z1p-lox; Site_1: Sal; Site_2: Not;
Lambda PRL2 is a cDNA library derived from equal
quantities of 4 pools of mRNA. The mRNA sources were 1) 7
day germinated etiolated seedlings; 2) tissue culture
grown roots; 3) staged plants half with 24 hour light
cycle, half on 16 hr light, 8 hour dark- rosettes; 4)
same plants as 3 but aerial tissue (stems, flowers and
siliques. The vector is BRL's lambda Z1p-lox. The cDNA
inserts were directionally cloned with Sal-Not arms using
oligo dt primed cDNA."
Site_1: 40 c
Site_2: 82 g
Site_3: 82 t
BASE COUNT 97 a 40 c 82 g . 82 t 5 others
ORIGIN

Alignment Scores:
Pred. No.: 515 Length: 306
Score: 53.00 Matches: 10
Percent Similarity: 81.258 Conservative: 3
Best Local Similarity: 62.508 Mismatches: 3
Query Match: 70.678 Indels: 0
DB: AV817296 Gaps: 0

US-10-075-846-4_COPY_192_207 (1-16) x AA597350 (1-306)
QY
1 SerSettleucySerProleuProSerleuSerleuServal 16
||||| ||| |||||||||||||||||||||:
Db
79 TCTATTACAGTAGCTTCTCCACCTCTCCCTCAATCTCTGATCACTT 32

RESULT 8
AV817296 404 bp mRNA linear EST 01-APR-2002
LOCUS AV817296
DEFINITION AV817296 RAFL9 Arabidopsis thaliana cDNA clone RAF109-93-O12 3',
mRNA sequence.
ACCESSION AV817296
VERSION AV817296.1 GI:19859167
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 404)
Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,
```





```

US-10-075-846-4_COPY_192_207 (1-16) x AV548451 (1-547)
Oy 1 SerSerSerIleuCySerProleuProSerIeuSerIeuSerVal 16
    ||| |||:: ||| ||||| |||||::|||::||| |||
    359 TCTATTTCAGTACGCTGTGCCACCTCTCCCTGCATCTCTGTATAGTT 406

RESULT 11
LOCUS AV548438 548 bp mRNA linear EST 06-SEP-2000
DEFINITION AV548438 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
ACCESSION CDNA Clone RFL54C06F 3', mRNA sequence.
VERSION AV548438.1 GI:8719851
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 548)
Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
DNA Res. 7, 175-180 (2000)
20363093

JOURNAL Contact: Erika Asamizu
MEDLINE The First Laboratory for Plant Gene Research
COMMENT Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES
     source
         1..548
             /organism="Arabidopsis thaliana"
             /strain="Columbia"
             /db_xref="taxon:3702"
             /clone="RZL54C06F"
             /clone_1lb="Arabidopsis thaliana roots Columbia"
             /tissue_type="roots"
             /note="Vector: pluescriptII SK-; Site_1: EcoRI; Site_2:
             XhoI"

BASE COUNT 176 a 131 c 84 g 157 t

ORIGIN

Alignment Scores:
Pred. No.: 1.17e+03 Length: 547
Score: 53.00 Matches: 10
Percent Similarity: 81.25% Conservative: 3
Best Local Similarity: 62.50% Mismatches: 3
Query Match: 70.67% Indels: 0
DB: 10 Gaps: 0

US-10-075-846-4_COPY_192_207 (1-16) x AV548451 (1-547)
Oy 1 SerSerSerIleuCySerProleuProSerIeuSerIeuSerVal 16
    ||| |||:: ||| ||||| |||||::|||::||| |||
    359 TCTATTTCAGTACGCTGTGCCACCTCTCCCTGCATCTCTGTATAGTT 406

RESULT 11
LOCUS AV548438 548 bp mRNA linear EST 06-SEP-2000
DEFINITION AV548438 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
ACCESSION CDNA Clone RFL54C06F 3', mRNA sequence.
VERSION AV548438.1 GI:8719851
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 548)
Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
DNA Res. 7, 175-180 (2000)
20363093

JOURNAL Contact: Erika Asamizu
MEDLINE The First Laboratory for Plant Gene Research
COMMENT Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES
     source
         1..548
             /organism="Arabidopsis thaliana"
             /strain="Columbia"
             /db_xref="taxon:3702"
             /clone="RZL54C06F"
             /clone_1lb="Arabidopsis thaliana roots Columbia"
             /tissue_type="roots"
             /note="Vector: pluescriptII SK-; Site_1: EcoRI; Site_2:
             XhoI"

BASE COUNT 176 a 131 c 84 g 157 t

ORIGIN

Alignment Scores:
Pred. No.: 1.17e+03 Length: 547
Score: 53.00 Matches: 10
Percent Similarity: 81.25% Conservative: 3
Best Local Similarity: 62.50% Mismatches: 3
Query Match: 70.67% Indels: 0
DB: 10 Gaps: 0

US-10-075-846-4_COPY_192_207 (1-16) x AV548451 (1-547)
Oy 1 SerSerSerIleuCySerProleuProSerIeuSerIeuSerVal 16
    ||| |||:: ||| ||||| |||||::|||::||| |||
    359 TCTATTTCAGTACGCTGTGCCACCTCTCCCTGCATCTCTGTATAGTT 406

RESULT 11
LOCUS AV548438 548 bp mRNA linear EST 06-SEP-2000
DEFINITION AV548438 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
ACCESSION CDNA Clone RFL54C06F 3', mRNA sequence.
VERSION AV548438.1 GI:8719851
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 548)
Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
DNA Res. 7, 175-180 (2000)
20363093

JOURNAL Contact: Erika Asamizu
MEDLINE The First Laboratory for Plant Gene Research
COMMENT Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES
     source
         1..548
             /organism="Arabidopsis thaliana"
             /strain="Columbia"
             /db_xref="taxon:3702"
             /clone="RZL54C06F"
             /clone_1lb="Arabidopsis thaliana roots Columbia"
             /tissue_type="roots"
             /note="Vector: pluescriptII SK-; Site_1: EcoRI; Site_2:
             XhoI"

BASE COUNT 176 a 131 c 84 g 157 t

ORIGIN

Alignment Scores:
Pred. No.: 1.17e+03 Length: 547
Score: 53.00 Matches: 10
Percent Similarity: 81.25% Conservative: 3
Best Local Similarity: 62.50% Mismatches: 3
Query Match: 70.67% Indels: 0
DB: 10 Gaps: 0

US-10-075-846-4_COPY_192_207 (1-16) x AV548451 (1-547)
Oy 1 SerSerSerIleuCySerProleuProSerIeuSerIeuSerVal 16
    ||| |||:: ||| ||||| |||||::|||::||| |||
    359 TCTATTTCAGTACGCTGTGCCACCTCTCCCTGCATCTCTGTATAGTT 406

RESULT 11
LOCUS AV548438 548 bp mRNA linear EST 06-SEP-2000
DEFINITION AV548438 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
ACCESSION CDNA Clone RFL54C06F 3', mRNA sequence.
VERSION AV548438.1 GI:8719851
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 548)
Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
DNA Res. 7, 175-180 (2000)
20363093

JOURNAL Contact: Erika Asamizu
MEDLINE The First Laboratory for Plant Gene Research
COMMENT Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES
     source
         1..548
             /organism="Arabidopsis thaliana"
             /strain="Columbia"
             /db_xref="taxon:3702"
             /clone="RZL54C06F"
             /clone_1lb="Arabidopsis thaliana roots Columbia"
             /tissue_type="roots"
             /note="Vector: pluescriptII SK-; Site_1: EcoRI; Site_2:
             XhoI"

BASE COUNT 176 a 131 c 84 g 157 t

ORIGIN

Alignment Scores:
Pred. No.: 1.17e+03 Length: 547
Score: 53.00 Matches: 10
Percent Similarity: 81.25% Conservative: 3
Best Local Similarity: 62.50% Mismatches: 3
Query Match: 70.67% Indels: 0
DB: 10 Gaps: 0

US-10-075-846-4_COPY_192_207 (1-16) x AV548451 (1-547)
Oy 1 SerSerSerIleuCySerProleuProSerIeuSerIeuSerVal 16
    ||| |||:: ||| ||||| |||||::|||::||| |||
    359 TCTATTTCAGTACGCTGTGCCACCTCTCCCTGCATCTCTGTATAGTT 406

RESULT 11
LOCUS AV548438 548 bp mRNA linear EST 06-SEP-2000
DEFINITION AV548438 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
ACCESSION CDNA Clone RFL54C06F 3', mRNA sequence.
VERSION AV548438.1 GI:8719851
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 548)
Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
DNA Res. 7, 175-180 (2000)
20363093

JOURNAL Contact: Erika Asamizu
MEDLINE The First Laboratory for Plant Gene Research
COMMENT Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES
     source
         1..548
             /organism="Arabidopsis thaliana"
             /strain="Columbia"
             /db_xref="taxon:3702"
             /clone="RZL54C06F"
             /clone_1lb="Arabidopsis thaliana roots Columbia"
             /tissue_type="roots"
             /note="Vector: pluescriptII SK-; Site_1: EcoRI; Site_2:
             XhoI"

BASE COUNT 176 a 131 c 84 g 157 t

ORIGIN

Alignment Scores:
Pred. No.: 1.17e+03 Length: 547
Score: 53.00 Matches: 10
Percent Similarity: 81.25% Conservative: 3
Best Local Similarity: 62.50% Mismatches: 3
Query Match: 70.67% Indels: 0
DB: 10 Gaps: 0

US-10-075-846-4_COPY_192_207 (1-16) x AV548451 (1-547)
Oy 1 SerSerSerIleuCySerProleuProSerIeuSerIeuSerVal 16
    ||| |||:: ||| ||||| |||||::|||::||| |||
    359 TCTATTTCAGTACGCTGTGCCACCTCTCCCTGCATCTCTGTATAGTT 406

RESULT 11
LOCUS AV548438 548 bp mRNA linear EST 06-SEP-2000
DEFINITION AV548438 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
ACCESSION CDNA Clone RFL54C06F 3', mRNA sequence.
VERSION AV548438.1 GI:8719851
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

```

```

DB          396  TCTATTACGAGCTTGTCACACCTCTCCCTCCATCTCTGTATCATGTT 443
RESULT 12
LOCUS      AV547337                      556 bp  mRNA  Linear  EST 06-SEP-2000
DEFINITION Arabidopsis thaliana roots Columbia Arabidopsis thaliana
ACCESSION  AV547337
VERSION    AV547337.1  GI:8718751
KEYWORDS
SOURCE     thale cress.
ORGANISM   Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE  1 (bases 1 to 556)
AUTHORS    Asamizu E., Nakamura Y., Sato S. and Tabata S.
TITLE      A large scale analysis of cDNA in Arabidopsis thaliana: Generation
            of 12,028 non-redundant expressed sequence tags from normalized and
            size-selected cDNA libraries
            DNA Res. 7, 175-180 (2000)
COMMENT    Contact: Erika Asamizu
            The First Laboratory for Plant Gene Research
            Kazusa DNA Research Institute
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
            Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES
            source
                1..556
                /organism="Arabidopsis thaliana"
                /strain="Columbia"
                /db_xref="taxon:3702"
                /clone="RZL29h10F"
                /clone_lib="Arabidopsis thaliana roots Columbia"
                /tissue_type="roots"
                /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
                XhoI"
BASE COUNT  145 a 144 c 88 g 179 t
ORIGIN
Alignment Scores:
Pred. No.: 1.2e+03 Length: 556
Score: 53.00 Matches: 10
Best Local Similarity: 81.25% Conservative: 3
Percent Similarity: 62.50% Mismatches: 3
Query Match: 70.67% Indels: 0
DB: 10 Gaps: 0
US-10-075-846-4_COPY_192_207 (1-16) x AV547337 (1-556)
Oy 1 1serSer1leLeuCySerProLeuProSerLeuSerLeuSerVal 16
    ||| ||||: ||| ||||| |||||: |||||: |||||
Ddb 221 TCTATTACGAGCTTGTCACACCTCTCCCTCCATCTCTGTATCATGTT 268

RESULT 13
LOCUS      AQ415379                      597 bp  DNA  Linear  GSS 23-MAR-1999
DEFINITION Arabidopsis thaliana roots Columbia Arabidopsis thaliana
ACCESSION  AQ415379
VERSION    AQ415379.1  GI:4474348
KEYWORDS   GSS.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE  1 (bases 1 to 597)
AUTHORS    Zhao S., Adams M.D., Nierman W., Malek J., de Jong P. and Venter
            J.C.
TITLE      Use of BAC End Sequences from Library RPCR-11 for Sequence-Ready
            Map Building
            Unpublished (1997)

```

COMMENT

Other:GSSS: RPCT-11-177K5.TV  
Contact: Shaying Zhao, William Nierman, Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: hbe@tigr.org  
Clones are derived from the human BAC library RPCT-11. For BAC library availability, please contact Pieter de Jong (pleter@e.jong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from Research Genet cs ([http://www.tigr.org/tbdb/humgen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tigr.org/tbdb/humgen/bac_end_search/bac_end_search.html)).  
Seq primer: SP6  
Class: BAC ends

FEATURES

source

1..597  
/organism="Homo sapiens"  
/db\_xref="GDB:7567828".  
/db\_xref="taxon:9606"  
/clone="RPCT-11-177K5"  
/clone\_lib="RPCT-11"  
/sex="Male"  
/cell\_type="Lymphocytes"  
/note="Vector: pBAC3.6; Site\_1: EcoRI; Site\_2: EcoRI; RPCT11 Human Male BAC Library"

BASE COUNT 142 a 163 c 127 g 164 t 1 others

ORIGIN

Alignment Scores:  
Pred. No.: 1.32e+03 Length: 597  
Score: 53.00 Matches: 9  
Percent Similarity: 92.86% Conservative: 4  
Best Local Similarity: 64.29% Mismatches: 1  
Query Match: 70.67% Indels: 0  
DB: 17 Gaps: 0

US-10-075-846-4\_COPY\_192-207 (1-16) x A0415379 (1-597)

QY 2 SerSerIleuCySerProIeuProSerIeuSer 15  
|||||:||||| |||||:|||||:|||||:|||||

Db 105 AGTCTTGCATGTGGCCCTCTCCCAACATGGCATTTC 146

RESULT 14  
B0246138 600 bp mRNA linear EST 03-MAY-2002  
LOCUS B0246138/c  
DEFINITION TaEI5015F10R TaEI5 Triticum aestivum cDNA clone TaEI5015F10R, mRNA  
sequence.  
ACCESSION B0246138  
VERSION B0246138.1 GI:20442014  
KEYWORDS EST.  
SOURCE bread wheat.  
ORGANISM Triticum aestivum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae  
; Triticeae; Triticum.  
1 (bases 1 to 600)  
Cloutier,S.  
Wheat functional genomics - Glanlea developing seeds cDNA libraries  
Unpublished (2002)  
Contact: Dr. Sylvie Cloutier  
Cereals Research Centre, Agriculture and Agri-food Canada  
155 Dufferin Rd, Winnipeg, MB, Canada R3T 2M9  
Tel: (204) 983-2340  
Fax: (204) 983-4604  
Email: scloutier@em.agr.ca  
was cloned directionally, not all sequences generated with reverse  
primer were from the 5' end (same with forward primer and 3' end).  
Average insert size 1s >1.4 kb  
Plate: 015 Row: F Column: 10  
Seq primer: M13 Reverse.  
Location/Qualifiers

FEATURES

source

1. .600

/organism="Triticum aestivum"

/cultivar="Glenlea"

/db\_xref="taxon:4565"

/clone="TAEL15015F10R"

/clone\_1id="TAEL15"

/tissue\_type="developing seeds"

/dev\_stage="15 days after anthesis"

/lab\_host="E. coli DH10s"

/note="Vector: pCMV-SPORT6.0 (Invitrogen Technologies); Site\_1: NotI; Site\_2: MluI; mRNA obtained from wheat seeds of cultivar Glenlea 15 days post-anthesis"

BASE COUNT 185 a 126 c 183 g 106 t

ORIGIN

Alignment Scores:

Prod. NO.:

Score: 1.33e+03 Length: 600

Percent Similarity: 53.00 Matches: 11

Best Local Similarity: 75.00% Conservative: 1

Query Match: 68.75% Mismatches: 4

DB: 70.67% Indels: 0

Gaps: 14

US-10-075-846-4\_COPY\_192\_207 (1-16) x BQ246138 (1-600)

Oy 1 SerSerSerIleLeuCySerProLeuProSerLeuSerLeuSerVal 16

Db 67 AGCAGCTCTTCTGCTGCTGACCCCTCTGCTGCTTGTGACACAGCCTT 20

RESULT 15

LOCUS BJ268516/c 615 bp mRNA linear EST 09-APR-2002

DEFINITION BJ268516 Y. Ogihara unpublished cDNA library, wh\_oh Triticum aestivum cDNA clone whoh17f17 5', mRNA sequence.

ACCESSION BJ268516

VERSION BJ268516.1 GI:20088515

KEYWORDS EST.

SOURCE bread wheat.

ORGANISM Triticum aestivum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae

REFERENCE 1 (bases 1 to 615)

AUTHORS Ogihara,Y. and Murai,K.

TITLE Expressed genes in Triticum aestivum

JOURNAL Unpublished (2002)

COMMENT Contact: Tadasu Shin-i

Center For Genetic Resource Information.

National Institute of Genetics

111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshini@genes.nig.ac.jp.

FEATURES

Source

1. .615

Location/Qualifiers

/organism="Triticum aestivum"

/cultivar="Chinese Spring"

/db\_xref="taxon:4565"

/clone="whoh17f17"

/clone\_1id="Y. Ogihara unpublished cDNA library, wh\_oh"

/tissue\_type="pistil at heading date"

/dev\_stage="Feekes" scale 10.5"

/note="Vector: Lambda Uni-ZAP XR, excised phagemid; Site\_1: EcoRI; Site\_2: XhoI; Plants were grown under hydroponic conditions at UC Davis, salt stressed for 12 hours, and for 7 days, then dissected and frozen (Akhunov in J. Dvorak Lab). Total RNA was prepared from sheath tissue, equal quantities of RNA were pooled from the two samples, polyA was purified from the pooled RNA, a cDNA library was made, and the cDNA clones were in vivo excised to give Bluescript phagemids in the T3 Close lab at the University of California, Riverside (Akhunov, Chai, Choi, Close, Fenton, Kianian, Otto, Simons, Zhang).

BASE COUNT	ORIGIN
184 a	127 c
	195 g
	108 t
	1 others

BASE COUNT	184 a	127 c	195 g	108 t	1 others
ORIGIN					

Alignment Scores:

Pred. No.:	1.38e+03	length:	615
Score:	53.00	Matches:	11
Percent Similarity:	75.00%	Conservative:	1
Best Local Similarity:	68.75%	Mismatches:	4
Query Match:	70.67%	Indels:	0
DB:	13	Gaps:	0

Score:	53.00	Matches:	11
Percent correct:	75.00%		

Percent Similarity:	75.00%	Conservative:	1
Best Local Similarity:	68.75%	Mismatches:	4

Query Match:	70.678	Indels:	0
--------------	--------	---------	---

DB: 13 Gaps: 0

DB: 13 Gaps: 0

US-10-075-846-4\_COPY\_192\_207 (1-16) x BJ268516 (1-615)

```

QY      1 SerSerSerIleLeuCysSerProLeuProSerLeuSerLeuSerVal 16
          |||||      |||||      |||||      |||::
Db      189 AGCAGCTCTTCTGCTGTACCCCTCTCCCTGGTTTGAGCAGCAGCTT 1422

```

```

Db      189 AGCAGCTTTCTGCTGTAGCCCTCTCCCTGGTTTGAGCAGCAGCCCTT 1422
          |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

Search completed: July 1, 2003, 00:55:32  
Job time : 89.1163 secs

Job time : 89.1163 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 25, 2003, 17:12:14 ; Search time 2.32662 Seconds  
(without alignments)  
916.354 Million cell updates/sec

Title: US-10-075-846-4\_COPY\_192\_207  
Perfect score: 75  
Sequence: 1 SSSIICSPILPSLSLV 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues  
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_10102.\*

1:	/SID52/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
2:	/SID52/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
3:	/SID52/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
4:	/SID52/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
5:	/SID52/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
6:	/SID52/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
7:	/SID52/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
8:	/SID52/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
9:	/SID52/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
10:	/SID52/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
11:	/SID52/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
12:	/SID52/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
13:	/SID52/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
14:	/SID52/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
15:	/SID52/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
16:	/SID52/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
17:	/SID52/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
18:	/SID52/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
19:	/SID52/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
20:	/SID52/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
21:	/SID52/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
22:	/SID52/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*
23:	/SID52/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	65.3	826	21	AA38474
2	49	65.3	855	21	AA38473
3	49	65.3	888	21	AA38472
4	47	62.7	83	22	AA694754
5	47	62.7	160	22	AA695892
6	46	61.3	490	23	AB82307
7	45.5	60.7	67	22	AB823990
8	45.5	60.7	89	22	AB811062
9	45	60.0	166	22	AA23635
10	45	60.0	166	22	AA14387

11	44	58.7	75	20	AAV14488	Fragment of human
12	44	58.7	120	22	AB817174	Human nervous syst
13	44	58.7	149	21	AA600176	Human secreted pro
14	44	58.7	887	22	AA39431	Human polypeptide
15	43	57.3	274	22	AA39109	Human polypeptide
16	43	57.3	289	22	AA40895	Human polypeptide
17	43	57.3	3594	23	AAE20147	Mouse C3b/C4b comp
18	42	56.0	39	22	AB815852	Human nervous syst
19	42	56.0	739	19	AA661532	Mouse fas-binding
20	42	56.0	739	22	AA669150	Mouse dact protein
21	41.5	55.3	73	22	AA869500	Human Immune/Haema
22	41	54.7	50	22	AA60466	Proionibacterium
23	41	54.7	69	22	AA602995	Proionibacterium
24	41	54.7	70	23	ABP05968	Human ORFX protein
25	41	54.7	73	22	AA59452	Proionibacterium
26	41	54.7	109	22	AA89167	Human Immune/Haema
27	41	54.7	110	22	AAU44557	Proionibacterium
28	41	54.7	216	22	AA698446	Gorilla olfactory
29	41	54.7	223	20	AAV74114	Human prostate tum
30	41	54.7	538	23	AAO21708	Human secreted pro
31	40	53.3	60	22	AAU33291	Novel human secret
32	40	53.3	68	23	ABP03420	Human OREX protein
33	40	53.3	72	22	AAU46650	Proionibacterium
34	40	53.3	93	22	AAU49188	Proionibacterium
35	40	53.3	98	21	AAV73448	Human secreted pro
36	40	53.3	106	21	AA640561	Human ORFX ORF35
37	40	53.3	133	23	AAO17460	Human Liver cancer
38	40	53.3	141	21	AAV54348	Amino acid sequenc
39	40	53.3	209	22	ABG04149	Novel human diagno
40	40	53.3	262	20	AAV37102	Chlamydia trachoma
41	40	53.3	265	22	AAU16053	Human novel secret
42	40	53.3	296	21	AA624854	Arabidopsis thalia
43	40	53.3	301	21	AA624853	Arabidopsis thalia
44	40	53.3	353	22	AB816162	Novel human diagno
45	40	53.3	395	21	AA642022	Arabidopsis thalia

## ALIGNMENTS

RESULT 1	AA38474	standard; Protein: 826 AA.
ID	AA38474	
XX	AA38474:	
DT	18-OCT-2000	(first entry)
XX		
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 47470.	
XX		
KW	Protein identification: signal transduction pathway; metabolic pathway;	
KW	hybridisation assay; genetic mapping; gene expression control; promoter;	
KW	termination sequence.	
XX		
OS	Arabidopsis thaliana.	
XX		
PN	EP1033405-A2.	
PD	06-SEP-2000.	
XX		
PF	25-FEB-2000; 2000EP-0301439.	
XX		
PR	25-FEB-1999;	99US-0121825.
PR	05-MAR-1999;	99US-0123180.
PR	09-MAR-1999;	99US-0123548.
PR	23-MAR-1999;	99US-0125788.
PR	25-MAR-1999;	99US-0126264.
PR	29-MAR-1999;	99US-0126785.
PR	01-APR-1999;	99US-0127462.
PR	06-APR-1999;	99US-0128234.
PR	08-APR-1999;	99US-0128714.
PR	16-APR-1999;	99US-0129845.
PR	19-APR-1999;	99US-0130077.

PR 21-APR-1999; 990S-0130449.  
PR 23-APR-1999; 990S-0130510.  
PR 23-APR-1999; 990S-0130891.  
PR 28-APR-1999; 990S-0131449.  
PR 30-APR-1999; 990S-0132048.  
PR 30-APR-1999; 990S-0132407.  
PR 04-MAY-1999; 990S-0132484.  
PR 05-MAY-1999; 990S-0132485.  
PR 06-MAY-1999; 990S-0132486.  
PR 07-MAY-1999; 990S-0132863.  
PR 11-MAY-1999; 990S-0134256.  
PR 14-MAY-1999; 990S-0134218.  
PR 14-MAY-1999; 990S-0134219.  
PR 14-MAY-1999; 990S-0134221.  
PR 18-MAY-1999; 990S-0134370.  
PR 19-MAY-1999; 990S-0134768.  
PR 19-MAY-1999; 990S-0134941.  
PR 20-MAY-1999; 990S-0135124.  
PR 21-MAY-1999; 990S-0135353.  
PR 24-MAY-1999; 990S-0135629.  
PR 25-MAY-1999; 990S-0136021.  
PR 27-MAY-1999; 990S-0136392.  
PR 28-MAY-1999; 990S-0136782.  
PR 01-JUN-1999; 990S-0137222.  
PR 03-JUN-1999; 990S-0137528.  
PR 04-JUN-1999; 990S-0137502.  
PR 07-JUN-1999; 990S-0137724.  
PR 08-JUN-1999; 990S-0138094.  
PR 10-JUN-1999; 990S-0138540.  
PR 10-JUN-1999; 990S-0138847.  
PR 14-JUN-1999; 990S-0139119.  
PR 16-JUN-1999; 990S-0139453.  
PR 16-JUN-1999; 990S-0139452.  
PR 17-JUN-1999; 990S-0139454.  
PR 18-JUN-1999; 990S-0139454.  
PR 18-JUN-1999; 990S-0139455.  
PR 18-JUN-1999; 990S-0139456.  
PR 18-JUN-1999; 990S-0139457.  
PR 18-JUN-1999; 990S-0139458.  
PR 18-JUN-1999; 990S-0139459.  
PR 18-JUN-1999; 990S-0139460.  
PR 18-JUN-1999; 990S-0139461.  
PR 18-JUN-1999; 990S-0139462.  
PR 18-JUN-1999; 990S-0139463.  
PR 18-JUN-1999; 990S-0139750.  
PR 18-JUN-1999; 990S-0139763.  
PR 21-JUN-1999; 990S-0139817.  
PR 22-JUN-1999; 990S-0139899.  
PR 23-JUN-1999; 990S-0140353.  
PR 23-JUN-1999; 990S-0140354.  
PR 24-JUN-1999; 990S-0140695.  
PR 28-JUN-1999; 990S-0140823.  
PR 29-JUN-1999; 990S-0140991.  
PR 30-JUN-1999; 990S-0141287.  
PR 01-JUL-1999; 990S-0141642.  
PR 02-JUL-1999; 990S-0142154.  
PR 06-JUL-1999; 990S-0142055.  
PR 08-JUL-1999; 990S-0142390.  
PR 08-JUL-1999; 990S-0142803.  
PR 09-JUL-1999; 990S-0142920.  
PR 12-JUL-1999; 990S-0142977.  
PR 13-JUL-1999; 990S-0143542.  
PR 14-JUL-1999; 990S-0143624.  
PR 15-JUL-1999; 990S-0144005.  
PR 16-JUL-1999; 990S-0144085.  
PR 16-JUL-1999; 990S-0144086.  
PR 19-JUL-1999; 990S-0144325.  
PR 19-JUL-1999; 990S-0144331.  
PR 19-JUL-1999; 990S-0144332.  
PR 19-JUL-1999; 990S-0144333.  
PR 19-JUL-1999; 990S-0144334.  
PR 19-JUL-1999; 990S-0144335.  
  
PR 20-JUL-1999; 990S-0144352.  
PR 20-JUL-1999; 990S-0144632.  
PR 20-JUL-1999; 990S-0144684.  
PR 21-JUL-1999; 990S-0144814.  
PR 21-JUL-1999; 990S-0145086.  
PR 21-JUL-1999; 990S-0145088.  
PR 22-JUL-1999; 990S-0145087.  
PR 22-JUL-1999; 990S-0145087.  
PR 22-JUL-1999; 990S-0145089.  
PR 22-JUL-1999; 990S-0145192.  
PR 23-JUL-1999; 990S-0145145.  
PR 23-JUL-1999; 990S-0145218.  
PR 23-JUL-1999; 990S-0145224.  
PR 26-JUL-1999; 990S-0145276.  
PR 27-JUL-1999; 990S-0145913.  
PR 27-JUL-1999; 990S-0145918.  
PR 27-JUL-1999; 990S-0145919.  
PR 28-JUL-1999; 990S-0145951.  
PR 02-AUG-1999; 990S-0146386.  
PR 02-AUG-1999; 990S-0146388.  
PR 02-AUG-1999; 990S-0146389.  
PR 04-AUG-1999; 990S-0147038.  
PR 04-AUG-1999; 990S-0147204.  
PR 04-AUG-1999; 990S-0147302.  
PR 05-AUG-1999; 990S-0147192.  
PR 05-AUG-1999; 990S-0147260.  
PR 06-AUG-1999; 990S-0147303.  
PR 06-AUG-1999; 990S-0147416.  
PR 09-AUG-1999; 990S-0147493.  
PR 09-AUG-1999; 990S-0147935.  
PR 10-AUG-1999; 990S-0148171.  
PR 11-AUG-1999; 990S-0148319.  
PR 12-AUG-1999; 990S-0148341.  
PR 13-AUG-1999; 990S-0148565.  
PR 13-AUG-1999; 990S-0148684.  
PR 16-AUG-1999; 990S-0149368.  
PR 17-AUG-1999; 990S-0149175.  
PR 18-AUG-1999; 990S-0149426.  
PR 20-AUG-1999; 990S-0149722.  
PR 20-AUG-1999; 990S-0149723.  
PR 20-AUG-1999; 990S-0149929.  
PR 23-AUG-1999; 990S-0149902.  
PR 23-AUG-1999; 990S-0149930.  
PR 25-AUG-1999; 990S-0150566.  
PR 26-AUG-1999; 990S-0150884.  
PR 27-AUG-1999; 990S-0151065.  
PR 27-AUG-1999; 990S-0151066.  
PR 27-AUG-1999; 990S-0151080.  
PR 30-AUG-1999; 990S-0151080.  
PR 31-AUG-1999; 990S-0151303.  
PR 01-SEP-1999; 990S-0151930.  
PR 07-SEP-1999; 990S-0152363.  
PR 10-SEP-1999; 990S-0153070.  
PR 13-SEP-1999; 990S-0153758.  
PR 13-SEP-1999; 990S-0154018.  
PR 16-SEP-1999; 990S-0154039.  
PR 20-SEP-1999; 990S-0154779.  
PR 22-SEP-1999; 990S-0155139.  
PR 23-SEP-1999; 990S-0155486.  
PR 24-SEP-1999; 990S-0155659.  
PR 28-SEP-1999; 990S-0156458.  
PR 29-SEP-1999; 990S-0156596.  
PR 04-OCT-1999; 990S-0157117.  
PR 05-OCT-1999; 990S-0157753.  
PR 06-OCT-1999; 990S-0157865.  
PR 07-OCT-1999; 990S-0158029.  
PR 08-OCT-1999; 990S-0158232.  
PR 12-OCT-1999; 990S-0158369.  
PR 13-OCT-1999; 990S-0159293.  
PR 13-OCT-1999; 990S-0159294.  
PR 13-OCT-1999; 990S-0159295.  
PR 14-OCT-1999; 990S-0159329.  
PR 14-OCT-1999; 990S-0159330.

PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160960.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 65.3%; Score 49; DB 21; Length 826;  
Best Local Similarity 73.3%; Pred. No.: 19;  
Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 SSSILCPLPSLSLS 15  
Db 431 SSSILMDPVPPLSLSS 445

## RESULT 2

AAG38473  
ID AAG38473 standard; Protein: 855 AA.

XX AAG38473;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 47469.

KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.

XX Arabidopsis thaliana.

OS EPI033405-A2.

XX 06-SEP-2000.

PD 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123160.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 23-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0127462.

PR 01-APR-1999; 99US-0128234.

PR 06-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144332.  
PR 20-JUL-1999; 99US-0144632.  
PR 21-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.





PR 14-MAY-1999;	99US-0134221.
PR 14-MAY-1999;	99US-0134370.
PR 18-MAY-1999;	99US-0134378.
PR 19-MAY-1999;	99US-0134941.
PR 20-MAY-1999;	99US-0135124.
PR 21-MAY-1999;	99US-0135353.
PR 24-MAY-1999;	99US-0135602.
PR 25-MAY-1999;	99US-0136021.
PR 27-MAY-1999;	99US-0136392.
PR 28-MAY-1999;	99US-0136782.
PR 01-JUN-1999;	99US-0137222.
PR 03-JUN-1999;	99US-0137528.
PR 04-JUN-1999;	99US-0137502.
PR 07-JUN-1999;	99US-0137724.
PR 08-JUN-1999;	99US-0138094.
PR 10-JUN-1999;	99US-0138540.
PR 10-JUN-1999;	99US-0138847.
PR 14-JUN-1999;	99US-0139119.
PR 16-JUN-1999;	99US-0139452.
PR 16-JUN-1999;	99US-0139453.
PR 17-JUN-1999;	99US-0139492.
PR 18-JUN-1999;	99US-0139454.
PR 18-JUN-1999;	99US-0139455.
PR 18-JUN-1999;	99US-0139456.
PR 18-JUN-1999;	99US-0139457.
PR 18-JUN-1999;	99US-0139458.
PR 18-JUN-1999;	99US-0139459.
PR 18-JUN-1999;	99US-0139460.
PR 18-JUN-1999;	99US-0139461.
PR 18-JUN-1999;	99US-0139462.
PR 18-JUN-1999;	99US-0139463.
PR 18-JUN-1999;	99US-0139750.
PR 18-JUN-1999;	99US-0139763.
PR 21-JUN-1999;	99US-0139817.
PR 22-JUN-1999;	99US-0139899.
PR 23-JUN-1999;	99US-0140353.
PR 23-JUN-1999;	99US-0140354.
PR 24-JUN-1999;	99US-0140695.
PR 28-JUN-1999;	99US-0140823.
PR 29-JUN-1999;	99US-0140991.
PR 30-JUN-1999;	99US-0141287.
PR 01-JUL-1999;	99US-0141842.
PR 01-JUL-1999;	99US-0142154.
PR 02-JUL-1999;	99US-0142055.
PR 06-JUL-1999;	99US-0142390.
PR 08-JUL-1999;	99US-0142803.
PR 09-JUL-1999;	99US-0142920.
PR 12-JUL-1999;	99US-0142977.
PR 13-JUL-1999;	99US-0143542.
PR 14-JUL-1999;	99US-0143624.
PR 15-JUL-1999;	99US-0144005.
PR 16-JUL-1999;	99US-0144085.
PR 16-JUL-1999;	99US-0144086.
PR 19-JUL-1999;	99US-0144325.
PR 19-JUL-1999;	99US-0144331.
PR 19-JUL-1999;	99US-0144332.
PR 19-JUL-1999;	99US-0144333.
PR 19-JUL-1999;	99US-0144334.
PR 19-JUL-1999;	99US-0144335.
PR 20-JUL-1999;	99US-0144352.
PR 20-JUL-1999;	99US-0144632.
PR 20-JUL-1999;	99US-0144684.
PR 21-JUL-1999;	99US-0144814.
PR 21-JUL-1999;	99US-0145086.
PR 21-JUL-1999;	99US-0145088.
PR 22-JUL-1999;	99US-0145085.
PR 22-JUL-1999;	99US-0145087.
PR 22-JUL-1999;	99US-0145089.
PR 23-JUL-1999;	99US-0145192.
PR 23-JUL-1999;	99US-0145218.
PR 23-JUL-1999;	99US-0145224.
PR 26-JUL-1999;	99US-0145226.
PR 27-JUL-1999;	99US-0145913.
PR 27-JUL-1999;	99US-0145918.
PR 28-JUL-1999;	99US-0145951.
PR 02-AUG-1999;	99US-0146386.
PR 02-AUG-1999;	99US-0146388.
PR 03-AUG-1999;	99US-0146389.
PR 04-AUG-1999;	99US-0147038.
PR 04-AUG-1999;	99US-0147204.
PR 05-AUG-1999;	99US-0147302.
PR 05-AUG-1999;	99US-0147192.
PR 06-AUG-1999;	99US-0147260.
PR 06-AUG-1999;	99US-0147303.
PR 09-AUG-1999;	99US-0147416.
PR 09-AUG-1999;	99US-0147935.
PR 10-AUG-1999;	99US-0148171.
PR 11-AUG-1999;	99US-0148319.
PR 12-AUG-1999;	99US-0148341.
PR 13-AUG-1999;	99US-0148565.
PR 13-AUG-1999;	99US-0148684.
PR 16-AUG-1999;	99US-0149368.
PR 17-AUG-1999;	99US-0149175.
PR 18-AUG-1999;	99US-0149426.
PR 20-AUG-1999;	99US-0149722.
PR 20-AUG-1999;	99US-0149723.
PR 20-AUG-1999;	99US-0149929.
PR 23-AUG-1999;	99US-0149902.
PR 23-AUG-1999;	99US-0149930.
PR 25-AUG-1999;	99US-0150566.
PR 26-AUG-1999;	99US-0150884.
PR 27-AUG-1999;	99US-0151065.
PR 27-AUG-1999;	99US-0151066.
PR 30-AUG-1999;	99US-0151080.
PR 31-AUG-1999;	99US-0151303.
PR 01-SEP-1999;	99US-0151438.
PR 07-SEP-1999;	99US-0151930.
PR 10-SEP-1999;	99US-0152363.
PR 13-SEP-1999;	99US-0153070.
PR 13-SEP-1999;	99US-0153758.
PR 15-SEP-1999;	99US-0154018.
PR 16-SEP-1999;	99US-0154039.
PR 20-SEP-1999;	99US-0154179.
PR 22-SEP-1999;	99US-0155139.
PR 23-SEP-1999;	99US-0155486.
PR 24-SEP-1999;	99US-0155659.
PR 28-SEP-1999;	99US-0156458.
PR 29-SEP-1999;	99US-0156596.
PR 04-OCT-1999;	99US-0157117.
PR 05-OCT-1999;	99US-0157753.
PR 06-OCT-1999;	99US-0157865.
PR 07-OCT-1999;	99US-0158029.
PR 08-OCT-1999;	99US-0158232.
PR 12-OCT-1999;	99US-0158369.
PR 13-OCT-1999;	99US-0159293.
PR 13-OCT-1999;	99US-0159294.
PR 13-OCT-1999;	99US-0159295.
PR 14-OCT-1999;	99US-0159329.
PR 14-OCT-1999;	99US-0159330.
PR 14-OCT-1999;	99US-0159331.
PR 14-OCT-1999;	99US-0159637.
PR 14-OCT-1999;	99US-0159638.
PR 18-OCT-1999;	99US-0159584.
PR 21-OCT-1999;	99US-0160741.
PR 21-OCT-1999;	99US-0160767.
PR 21-OCT-1999;	99US-0160768.
PR 21-OCT-1999;	99US-0160770.
PR 21-OCT-1999;	99US-0160814.
PR 21-OCT-1999;	99US-0160815.
PR 22-OCT-1999;	99US-0160980.
PR 22-OCT-1999;	99US-0160981.
PR 22-OCT-1999;	99US-0160989.
PR 25-OCT-1999;	99US-0161404.

[illegible]



XX WO200210210-A2.  
XX  
XX 07-FEB-2002.  
XX  
XX 28-AUG-2001; 2001WO-EP09892.  
XX  
XX 28-AUG-2001; 2001WO-EP09892.  
XX  
XX (FARB ) BAYER AG.  
XX  
XX Tietjen K, Weidler M;  
XX  
XX WPI: 2002-269010/31.  
XX  
XX  
XX Identifying plant target proteins for herbicidally active compounds,  
XX comprising aligning and comparing nucleic acid or amino acid sequences  
XX from plant with nucleic acid or amino acid sequences from non-plant  
XX organisms -  
XX  
XX Claim 5; SEQ ID NO 1418; 261pp + Sequence Listing; English.  
XX  
XX The invention relates to identifying target proteins  
XX (AB90750-AB94016) for herbicidally active compounds, comprising  
XX aligning and comparing nucleic acid or amino acid sequences from plant  
XX with nucleic acid or amino acid sequences from non-plant organisms using  
XX suitable search parameters, where plant sequences having an E-value  
XX greater by a factor of 3 than the E-value of most similar non-plant  
XX sequences are selected. The polypeptides or nucleic acids encoding them  
XX are useful for identifying modulators. The identified modulators are  
XX useful as herbicides.  
XX  
XX  
XX Sequence 490 AA:  
XX  
XX Query Match 61.3%; Score 46; DB 23; Length 490;  
XX Best Local Similarity 68.8%; Pred. No. 33;  
XX Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
XX  
XX QY 1 SSSIICSPLPSPSLV 16  
XX :|||:|||||:|  
XX Db 210 SSSLLSPPLPPSLPV 225  
XX  
XX  
XX RESULT 7  
XX ABG23990  
XX ID ABG23990 standard; Protein: 67 AA.  
XX  
XX AC ABG23990;  
XX  
XX DT 18-FEB-2002 (first entry)  
XX  
XX DE Novel human diagnostic protein #23981.  
XX  
XX KW Human: chromosome mapping; gene mapping; gene therapy; forensic;  
XX food supplement; medical imaging; diagnostic; genetic disorder.  
XX  
XX OS Homo sapiens.  
XX  
XX PN WO200175067-A2.  
XX  
XX PD 11-OCT-2001.  
XX  
XX PF 30-MAR-2001; 2001WO-US08631.  
XX  
XX PR 31-MAR-2000; 2000US-0540217.  
XX  
XX PR 23-AUG-2000; 2000US-0649167.  
XX  
XX PA (HYSE-) HYSEQ INC.  
XX  
XX PI Drmanac RT, Liu C, Tang YT;  
XX  
XX WPI: 2001-639362/73.  
XX  
XX N-PSDB; AAS8177.  
XX  
XX

XX New isolated polynucleotide and encoded polypeptides, useful in  
XX diagnostics, forensics, gene mapping, identification of mutations  
XX responsible for genetic disorders or other traits and to assess  
XX biodiversity -  
XX  
XX Claim 20; SEQ ID No 54349; 103pp; English.  
XX  
XX The invention relates to isolated polynucleotide (I) and  
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,  
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
XX and gene mapping, and in recombinant production of (II). The  
XX polynucleotides are also used in diagnostics as expressed sequence tags  
XX for identifying expressed genes. (I) is useful in gene therapy techniques  
XX to restore normal activity of (II) or to treat disease states involving  
XX (II). (II) is useful for generating antibodies against it, detecting or  
XX quantitating a polypeptide in tissue, as molecular weight markers and as  
XX a food supplement. (II) and its binding partners are useful in medical  
XX imaging of sites expressing (II). (I) and (II) are useful for treating  
XX disorders involving aberrant protein expression or biological activity.  
XX The polypeptide and polynucleotide sequences have applications in  
XX diagnostics, forensics, gene mapping, identification of mutations  
XX responsible for genetic disorders or other traits to assess biodiversity  
XX and to produce other types of data and products dependent on DNA and  
XX amino acid sequences. ABG00010-ABG30377 represent novel human  
XX diagnostic amino acid sequences of the invention.  
XX Note: The sequence data for this patent did not appear in the printed  
XX specification, but was obtained in electronic format directly from WIPO  
XX at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX  
XX Sequence 67 AA:  
XX  
XX Query Match 60.7%; Score 45.5; DB 22; Length 67;  
XX Best Local Similarity 78.6%; Pred. No. 4.9;  
XX Matches 11; Conservative 1; Mismatches 1; Indels 1; Gaps 1;  
XX  
XX QY 1 SSSIICSPLPSPSL 14  
XX :|||:|||||:|  
XX Db 20 TSSFLCPLP-LSL 32  
XX  
XX  
XX RESULT 8  
XX ABG11062  
XX ID ABG11062 standard; Protein: 89 AA.  
XX  
XX AC ABG11062;  
XX  
XX DT 18-FEB-2002 (first entry)  
XX  
XX DE Novel human diagnostic protein #11053.  
XX  
XX KW Human: chromosome mapping; gene mapping; gene therapy; forensic;  
XX food supplement; medical imaging; diagnostic; genetic disorder.  
XX  
XX OS Homo sapiens.  
XX  
XX PN WO200175067-A2.  
XX  
XX PD 11-OCT-2001.  
XX  
XX PF 30-MAR-2001; 2001WO-US08631.  
XX  
XX PR 31-MAR-2000; 2000US-0540217.  
XX  
XX PR 23-AUG-2000; 2000US-0649167.  
XX  
XX PA (HYSE-) HYSEQ INC.  
XX  
XX PI Drmanac RT, Liu C, Tang YT;  
XX  
XX WPI: 2001-639362/73.  
XX  
XX N-PSDB; AAS75249.  
XX  
XX New isolated polynucleotide and encoded polypeptides, useful in  
XX

PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity  
PS Claim 20; SEQ ID No 41421; 103bp; English.  
XX  
XX The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human  
CC diagnostic amino acid sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 89 AA:  
Query Match 60.7%; Score 45.5; DB 22; Length 89;  
Best Local Similarity 78.6%; Pred. No. 6.6;  
Matches 11; Conservative 1; Mismatches 1; Indels 1; Gaps 1;  
QY 1 SSITCSPLPSTSL 14  
DB 49 TSSFICSPLP-LSL 61  
RESULT 9  
AA023635  
ID AA023635 standard; Protein; 166 AA.  
AC AA023635;  
XX  
XX 12-OCT-2001 (first entry)  
DT  
XX  
XX HHV-2 EST encoded protein SEQ ID NO: 1160.  
DE  
XX Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;  
KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;  
KW diagnostics; forensic test; gene mapping; genetic disorder;  
KW biodiversity; gene therapy; nutrition.  
XX  
XX Human herpesvirus 2.  
OS  
XX  
XX WO200154477-A2.  
PN  
XX  
XX 02-AUG-2001.  
PD  
XX  
XX 25-JAN-2001; 2001WO-US02687.  
PF  
XX  
XX 25-JAN-2000; 2000US-0491404.  
PR 17-JUL-2000; 2000US-0617746.  
PR 03-AUG-2000; 2000US-0631451.  
PR 15-SEP-2000; 2000US-0663870.  
XX  
XX (HYSE-) HYSEQ INC.  
PA  
XX Tang YF, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;  
PI Cao Y, Drmanac RA, Zhang J, Werhman T;  
XX  
XX WPI: 2001-476164/51.  
DR

DR N-PSDB; AAH98294.  
XX  
XX Isolated polypeptide for treatment of diseases, diagnostics, raising  
PT antibodies and research use -  
XX  
XX Claim 20; Page 850-851; 1275bp; English.  
PS  
XX  
XX The present invention provides the protein and coding sequences of novel  
CC proteins from a variety of organisms, including human, dog, cat, horse,  
CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea  
CC urchin and tomato. These were derived from expressed sequence tags (ESTs)  
CC from the organism of interest. They can be used in diagnostics,  
CC forensics, gene mapping, identification of mutations, to assess  
CC biodiversity and for nutritional purposes. The present sequence is a  
CC protein of the invention.  
XX  
SQ Sequence 166 AA:  
Query Match 60.0%; Score 45; DB 22; Length 166;  
Best Local Similarity 66.7%; Pred. No. 15;  
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 5 LCSPPLPSLSLV 16  
DB 143 LCPPLPSLSCAI 154  
RESULT 10  
AA014387  
ID AA014387 standard; Protein; 166 AA.  
AC AA014387;  
XX  
XX 24-OCT-2001 (first entry)  
DT  
XX  
XX Human novel protein #258.  
DE  
XX  
XX Human; novel protein; antianaemic; osteopathic; antiinflammatory;  
KW immunomodulatory; cyostatic; neuroprotective; vulnerrary; nootropic;  
KW anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral;  
KW antibacterial; antiallergic; dermatological; hemostatic; antiasthmatic;  
KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;  
KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;  
KW tissue regeneration; immune disorder.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200154437-A2.  
PN  
XX  
XX 02-AUG-2001.  
PD  
XX  
XX 25-JAN-2001; 2001WO-US02623.  
PF  
XX  
XX 25-JAN-2000; 2000US-0491404.  
PR  
XX  
XX (HYSE-) HYSEQ INC.  
PA  
XX Tang YF, Liu C, Drmanac RT;  
PI  
XX  
XX WPI: 2001-451939/48.  
DR N-PSDB; AAS22692.  
XX  
XX Isolated polypeptides useful for treating anti-inflammatory diseases,  
PT nervous system disorders, and for regenerating bone and cartilage -  
XX  
XX Example 4; Page 795-796; 894bp; English.  
PS  
XX  
XX The invention relates to polynucleotides encoding novel human  
CC proteins or their active domains. The polypeptides, polynucleotides and  
CC antibodies raised against the polypeptides are used in a method of  
CC treatment of a mammal and prevention of disorders caused by the aberrant  
CC protein expression or activity. The polypeptides can be used as  
CC molecular weight markers, food supplements, and in antibody production.  
CC

CC The polypeptides are used to identify compounds which bind to the  
 CC polypeptides. Polynucleotides of the invention are used as probes and  
 CC primers, for sequencing, for chromosome or gene mapping, in the  
 CC production of recombinant proteins, and in generating anti-sense DNA or  
 CC RNA and in gene therapy. Polypeptides of the invention can be used to  
 CC target drugs to a tumour. In assays to determine biological activity, to  
 CC raise antibodies/ elicit an immune response, to determine quantitative  
 CC protein levels, as tissue markers, and to isolate receptors or ligands.  
 CC Polypeptides of the invention may also be useful in treating platelet  
 CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,  
 CC ligament and/or nerve tissue, wound healing, treating burns, promoting  
 CC the proliferation, differentiation and survival of stem cells, as a  
 CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,  
 CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral  
 CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or  
 CC fungal infection or from autoimmunity, cancer, allergy, asthma,  
 CC graft-versus-host disease, eczema, haemophilia, thrombosis,  
 CC anti-inflammatory diseases, nervous system disorders, and infection.  
 CC The present sequence represents a protein of the invention.

SQ Sequence 166 AA;

Query Match Best Local Similarity 60.0%; Score 45; DB 22; Length 166;  
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 LCSPLPSLSLV 16  
 11 ||||| : :  
 Db 143 LCSPLPSLSLCAI 154

RESULT 11  
 ID AAY14488  
 AAY14488 standard; Protein: 75 AA.

XX AAY14488;

DT 17-AUG-1999 (first entry)

XX Fragment of human secreted protein encoded by gene 33.

XX Human; secreted protein; fusion protein; gene therapy; protein therapy;  
 KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;  
 KW developmental abnormality; foetal deficiency; blood; allergy; renal;  
 KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;  
 KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;  
 KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;  
 KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;  
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.

XX Homo sapiens.

PN WO9919339-A1.

XX 22-APR-1999.

PF 08-OCT-1998; 98WO-US21142.

PR 09-OCT-1997; 97US-0071498.

PR 09-OCT-1997; 97US-0061463.

PR 09-OCT-1997; 97US-0061527.

PR 09-OCT-1997; 97US-0061529.

PR 09-OCT-1997; 97US-0061532.

PR 09-OCT-1997; 97US-0061536.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Brewer LA, Duan R, Ebner R, Ferrie AM, Florence C;

PI Florence KA, Greene JM, Olsen HS, Rosen CA, Ruben SM;

PI Young PE, Yu G;

DR WPI; 1999-277587/23.

PT New isolated human genes and the secreted polypeptides they encode  
 XX disclosure; Page 45; 226pp; English.

CC This sequence represents a fragment of a secreted human protein encoded  
 CC by the nucleic acid molecule detailed in the descriptor line. The gene  
 CC can be used to generate fusion proteins by linking to the gene to a  
 CC human immunoglobulin Fc portion (e.g. AAX79002) for increasing the  
 CC stability of the fused protein as compared to the human protein only.  
 CC The invention relates to 53 novel genes and their fragments (nucleic  
 CC acid sequences: AAX79011-X79064; amino acid sequences AAY14411-Y14464)  
 CC which are useful for preventing, treating or ameliorating medical  
 CC conditions e.g. by protein or gene therapy. Also, pathological  
 CC conditions can be diagnosed by determining the amount of the new  
 CC polypeptides in a sample or by determining the presence of mutations in  
 CC the new polynucleotides. Specific uses are described for each of the 53  
 CC polynucleotides, based on which tissues they are most highly expressed in  
 CC (see AAX79011 for described uses).

SQ Sequence 75 AA;

Query Match Best Local Similarity 58.7%; Score 44; DB 20; Length 75;  
 Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 SSILCSPLPSLS 13  
 : ||||| : :  
 Db 15 TSQICSPMSPLN 26

RESULT 12

ABBI7174  
 ID ABB17174 standard; Protein: 120 AA.

XX ABB17174;

DT 23-JAN-2002 (first entry)

XX Human nervous system related polypeptide SEQ ID NO 5831.

XX Human; nootropic; neuroprotective; cytoskeletal; dermatological; vitruclide;  
 KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;  
 KW antiparkinsonian; antiskilling; antianaemic; antiarthritis; cancer;  
 KW antihemematic; hepatotropic; cerebroprotective; antiinflammatory;  
 KW antidiabetic; antidiabetic; antitumor; anticonvulsant; antifungal;  
 KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;  
 KW neurological disease; infection; nephrotropic; gene therapy; vaccine.

XX Homo sapiens.

PN WO200159063-A2.

XX 16-AUG-2001.

PF 17-JAN-2001; 2001WO-US01334.

PR 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.

PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216647.

PR 07-JUL-2000; 2000US-0216880.

PR 11-JUL-2000; 2000US-0217487.

PR 11-JUL-2000; 2000US-0217496.

PR 14-JUL-2000; 2000US-0218290.

PR 26-JUL-2000; 2000US-0220963.

PR 26-JUL-2000; 2000US-0220964.  
 PR 14-AUG-2000; 2000US-0224518.  
 PR 14-AUG-2000; 2000US-0224519.  
 PR 14-AUG-2000; 2000US-0225213.  
 PR 14-AUG-2000; 2000US-0225214.  
 PR 14-AUG-2000; 2000US-0225266.  
 PR 14-AUG-2000; 2000US-0225267.  
 PR 14-AUG-2000; 2000US-0225268.  
 PR 14-AUG-2000; 2000US-0225270.  
 PR 14-AUG-2000; 2000US-0225447.  
 PR 14-AUG-2000; 2000US-0225757.  
 PR 14-AUG-2000; 2000US-0225758.  
 PR 14-AUG-2000; 2000US-0225759.  
 PR 18-AUG-2000; 2000US-0226279.  
 PR 22-AUG-2000; 2000US-0226681.  
 PR 22-AUG-2000; 2000US-0226868.  
 PR 22-AUG-2000; 2000US-0227182.  
 PR 23-AUG-2000; 2000US-0227009.  
 PR 30-AUG-2000; 2000US-0228924.  
 PR 01-SEP-2000; 2000US-0229287.  
 PR 01-SEP-2000; 2000US-0229343.  
 PR 01-SEP-2000; 2000US-0229344.  
 PR 01-SEP-2000; 2000US-0229345.  
 PR 03-SEP-2000; 2000US-0229509.  
 PR 06-SEP-2000; 2000US-0230437.  
 PR 06-SEP-2000; 2000US-0230438.  
 PR 08-SEP-2000; 2000US-0231242.  
 PR 08-SEP-2000; 2000US-0231243.  
 PR 08-SEP-2000; 2000US-0231244.  
 PR 08-SEP-2000; 2000US-0231413.  
 PR 08-SEP-2000; 2000US-0231414.  
 PR 08-SEP-2000; 2000US-0232080.  
 PR 08-SEP-2000; 2000US-0232081.  
 PR 12-SEP-2000; 2000US-0231966.  
 PR 14-SEP-2000; 2000US-0232397.  
 PR 14-SEP-2000; 2000US-0232398.  
 PR 14-SEP-2000; 2000US-0232399.  
 PR 14-SEP-2000; 2000US-0232400.  
 PR 14-SEP-2000; 2000US-0232401.  
 PR 14-SEP-2000; 2000US-0233063.  
 PR 14-SEP-2000; 2000US-0233064.  
 PR 14-SEP-2000; 2000US-0233065.  
 PR 21-SEP-2000; 2000US-0234223.  
 PR 21-SEP-2000; 2000US-0234274.  
 PR 25-SEP-2000; 2000US-0234997.  
 PR 25-SEP-2000; 2000US-0234998.  
 PR 26-SEP-2000; 2000US-0235484.  
 PR 27-SEP-2000; 2000US-0235834.  
 PR 27-SEP-2000; 2000US-0235836.  
 PR 29-SEP-2000; 2000US-0236327.  
 PR 29-SEP-2000; 2000US-0236367.  
 PR 29-SEP-2000; 2000US-0236368.  
 PR 29-SEP-2000; 2000US-0236369.  
 PR 29-SEP-2000; 2000US-0236370.  
 PR 02-OCT-2000; 2000US-0236802.  
 PR 02-OCT-2000; 2000US-0237037.  
 PR 02-OCT-2000; 2000US-0237038.  
 PR 02-OCT-2000; 2000US-0237039.  
 PR 02-OCT-2000; 2000US-0237040.  
 PR 13-OCT-2000; 2000US-0239935.  
 PR 13-OCT-2000; 2000US-0239937.  
 PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241808.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 20-OCT-2000; 2000US-0241826.  
 PR 20-OCT-2000; 2000US-0242221.  
 PR 01-NOV-2000; 2000US-0244617.  
 PR 08-NOV-2000; 2000US-0246474.  
 PR 08-NOV-2000; 2000US-0246475.

PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 08-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249224.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 01-DEC-2000; 2000US-0251160.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251866.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCT INC.

Rosen CA, Barash SC, Ruben SM;

WPI, 2001-541565/60.

N-PSDB; ABA13500.

Nucleic acids encoding 3224 human nervous system antigen polypeptides, useful for preventing, diagnosing and/or treating nervous system cancers and metastases -

Claim 11; SEQ ID NO 5831; 1701pp + Sequence Listing; English.

The invention relates to novel genes (ABA11004-ABA21534) and proteins (ABA14678-ABA18001) useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and

CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
CC and parasitic infections.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX  
SQ Sequence 120 AA;

Query Match 58.7%; Score 44; DB 22; Length 120;  
Best Local Similarity 66.7%; Pred. No. 16;  
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 SSSIICSPPLPSLS 15  
11 111111:1  
Db 15 SSVOLCEPLPSLAES 29

RESULT 13  
AAG00176  
ID AAG00176 standard; Protein; 149 AA.

XX  
AC AAG00176;

DT 06-OCT-2000 (first entry)

XX  
DE Human secreted protein, SEQ ID NO: 4257.

XX  
KM Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
KM gene therapy; chromosome mapping.

XX  
OS Homo sapiens.

XX  
PN EP1033401-A2.

XX  
PD 06-SEP-2000.

XX  
PF 21-FEB-2000; 2000EP-0200610.

XX  
PR 26-FEB-1999; 99US-0122487.

XX  
PA (GENSET) GENSET.

XX  
PI Dumas Milne Edwards J, Duclert A, Giordano J;

XX  
DR WPI; 2000-500381/45.

XX  
DR N-PSDB; AAC00182.

XX  
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for  
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -  
XX  
PS Claim 13; SEQ ID 4257; 71pp + CD-ROM; English.

XX  
CC The present sequence is a polypeptide encoded by one of a large number  
CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs  
CC were prepared from total human RNAs or polyA+ RNAs derived from 30  
CC different tissues. EST sequences usually correspond mainly to the 3'  
CC untranslated region (UTR) of the mRNA because they are often obtained  
CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for  
CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in  
CC those cases where longer cDNA sequences have been obtained, the full 5'  
CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'  
CC ends and can therefore be used to obtain full length cDNAs and genomic  
CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and  
CC chromosome mapping procedures. They are used to obtain upstream  
CC regulatory sequences and to design expression and secretion vectors.

XX  
SQ Sequence 149 AA;

Query Match 58.7%; Score 44; DB 21; Length 149;  
Best Local Similarity 69.2%; Pred. No. 20;  
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SSSIICSPPLPS 13  
11 111111:1  
Db 85 SSGIIRSPPLPSVS 97

RESULT 14  
AAM39431  
ID AAM39431 standard; Protein; 887 AA.

XX  
AC AAM39431;

DT 22-OCT-2001 (first entry)

XX  
DE Human polypeptide SEQ ID NO 2576.

XX  
KM Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KM peripheral nervous system; neuropathy; central nervous system; CNS;  
KM Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KM amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KM chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KM leukaemia.

XX  
OS Homo sapiens.

XX  
PN WO200153312-A1.

XX  
PD 26-JUL-2001.

XX  
PF 26-DEC-2000; 2000WO-US34263.

XX  
PR 21-JAN-2000; 2000US-0488725.

XX  
PR 25-APR-2000; 2000US-0552317.

XX  
PR 09-JUL-2000; 2000US-0598042.

XX  
PR 19-JUL-2000; 2000US-0620312.

XX  
PR 03-AUG-2000; 2000US-0653450.

XX  
PR 14-SEP-2000; 2000US-0662191.

XX  
PR 19-OCT-2000; 2000US-0693036.

XX  
PR 29-NOV-2000; 2000US-0727344.

XX  
PA (HYSE-) HYSEQ INC.

XX  
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Weinman T, Xu C, Xue AJ, Yang Y, Zhang J;  
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX  
DR WPI: 2001-442253/47.

XX  
DR N-PSDB; AAI58587.

XX  
PT Novel nucleic acids and polypeptides, useful for treating disorders  
PT such as central nervous system injuries -

XX  
PS Example 4; SEQ ID NO 2576; 10078pp; English.

XX  
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and  
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localized neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC activating/inhibiting activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders.

XX  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification.

XX  
SQ Sequence 887 AA;

Query Match 58.7%; Score 44; DB 22; Length 887;



Best Local Similarity 69.2%; Pred. No. 1.3e+02;  
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 SSSILCSPLPSLS 13  
|||:|||||:|  
Db 442 SSOITTSPLPSVS 454

## RESULT 15

AAM39109

ID AAM39109 standard; Protein: 274 AA.

AC AAM39109;

DT 22-OCT-2001 (first entry)

DE Human polypeptide SEQ ID NO 2254.

XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;

KW peripheral nervous system; neuropathy; central nervous system; CNS;

KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;

KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;

KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;

KW leukaemia.

XX Homo sapiens.

XX WO200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US34263.

XX 21-JAN-2000; 2000US-0488725.

XX 25-APR-2000; 2000US-0552317.

XX 09-JUL-2000; 2000US-0596042.

XX 19-JUL-2000; 2000US-0620312.

XX 03-AUG-2000; 2000US-0653450.

XX 14-SEP-2000; 2000US-0662191.

XX 19-OCT-2000; 2000US-0693036.

XX 29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.

XX Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

XX PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;

XX PI Zhao QA, Zhou P, Goodrich R, Dirmann RT;

XX WPI: 2001-442253/47.

XX N-PSDB: AA158265.

XX Novel nucleic acids and polypeptides, useful for treating disorders

XX PT such as central nervous system injuries.

XX Example 4; SEQ ID NO 2254; 10078pp; English.

XX The invention relates to human nucleic acids (AA157798-AA161369) and

XX the encoded polypeptides (AAM38642-AAM42213) with nootropic,

XX immunosuppressant and cytostatic activity. The polynucleotides are useful

XX in gene therapy. A composition containing a polypeptide or polynucleotide

XX of the invention may be used to treat diseases of the peripheral nervous

XX system, such as peripheral nervous injuries, peripheral neuropathy and

XX localised neuropathies and central nervous system diseases, such as

XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic

XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the

XX utilisation of the activities such as: immune system suppression,

XX activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic

XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,

XX assays for receptor activity, arthritis and inflammation, leukaemias and

XX C.N.S disorders.

XX Note: The sequence data for this patent did not form part of the printed

XX specification.

SQ Sequence 274 AA;

Query Match

57.3%; Score 43; DB 22; Length 274;

Best Local Similarity

71.4%; Pred. No. 54;

Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Oy 1 SSSILCSPLPSLSL 14

|||:|||||

Db 145 SPSLQCPPLPSLSL 158

Search completed: June 25, 2003, 17:15:33

Job time : 3.32662 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 25, 2003, 17:14:25 ; Search time 0.85906 Seconds  
(without alignments)  
548.002 Million cell updates/sec

Title: US-10-075-846-4\_COPY\_192\_207

Sequence: 1 SSILCSPPLSLSLV 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued\_Patents\_AA:\*

1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/Dackfilest1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	58.7	75	4	US-09-288-143-205
2	43	57.3	106	4	US-09-177-249-208
3	42	56.0	739	4	US-09-022-983-2
4	39	52.0	291	4	US-09-328-869-2
5	39	52.0	291	4	US-09-629-774A-2
6	38	50.7	13	1	US-08-299-187-13
7	38	50.7	13	5	PCT-US95-11114-13
8	38	50.7	251	1	US-07-956-700B-84
9	38	50.7	251	1	US-08-476-537-84
10	38	50.7	251	1	US-08-485-607-84
11	38	50.7	251	2	US-08-475-879-84
12	38	50.7	251	2	US-09-433-043B-84
13	38	50.7	460	1	US-08-476-008-50
14	38	50.7	460	1	US-08-306-063-50
15	38	50.7	460	1	US-08-833-485-50
16	38	50.7	460	4	US-09-137-440-50
17	38	50.7	621	3	US-08-604-789B-4
18	38	50.7	621	4	US-09-312-721A-4
19	38	50.7	667	4	US-09-094-557-1
20	38	50.7	703	4	US-09-433-043B-124
21	38	50.7	943	4	US-08-808-982-7
22	38	50.7	943	4	US-09-306-902A-7
23	37	49.3	423	4	US-08-955-713-2
24	36	48.0	96	4	US-08-679-493A-147
25	36	48.0	569	1	US-08-306-231-3
26	36	48.0	676	2	US-08-398-590A-40
27	36	48.0	676	4	US-08-894-997-40

28	36	48.0	896	2	US-08-640-389A-10	Sequence 10, Appl
29	36	48.0	896	4	US-08-618-957A-10	Sequence 10, Appl
30	36	48.0	898	2	US-08-693-697-36	Sequence 36, Appl
31	36	48.0	898	4	US-08-588-189-3	Sequence 3, Appl
32	36	48.0	906	2	US-08-640-389A-9	Sequence 9, Appl
33	36	48.0	906	4	US-08-618-957A-9	Sequence 9, Appl
34	36	48.0	908	2	US-08-693-697-33	Sequence 33, Appl
35	36	48.0	908	2	US-08-588-526-3	Sequence 3, Appl
36	36	48.0	958	2	US-08-640-389A-8	Sequence 8, Appl
37	36	48.0	958	4	US-08-618-957A-8	Sequence 8, Appl
38	36	48.0	960	1	US-08-353-886A-8	Sequence 8, Appl
39	36	48.0	960	2	US-08-588-190-3	Sequence 3, Appl
40	36	48.0	960	2	US-08-693-697-8	Sequence 8, Appl
41	36	48.0	960	2	US-08-640-389A-3	Sequence 3, Appl
42	36	48.0	960	3	US-08-693-696-8	Sequence 3, Appl
43	36	48.0	960	4	US-08-618-957A-3	Sequence 3, Appl
44	36	48.0	1165	2	US-08-640-389A-11	Sequence 11, Appl
45	36	48.0	1165	2	US-08-599-455B-4	Sequence 4, Appl

#### ALIGNMENTS

```
RESULT 1
US-09-288-143-205
; Sequence 205, Application US/09288143
; Patent No. 6433139
; GENERAL INFORMATION:
; APPLICANT: Brewer et al.
; TITLE OF INVENTION: 53 Human Secreted Proteins
; FILE REFERENCE: P2018P1
; CURRENT APPLICATION NUMBER: US/09/288,143
; EARLIER FILING DATE: 1999-04-08
; EARLIER APPLICATION NUMBER: PCT/US98/21142
; EARLIER FILING DATE: 1998-10-08
; EARLIER APPLICATION NUMBER: 60/061,463
; EARLIER FILING DATE: 1997-10-09
; EARLIER APPLICATION NUMBER: 60/061,529
; EARLIER FILING DATE: 1997-10-09
; EARLIER APPLICATION NUMBER: 60/071,498
; EARLIER FILING DATE: 1997-10-09
; EARLIER APPLICATION NUMBER: 60/061,527
; EARLIER FILING DATE: 1997-10-09
; EARLIER APPLICATION NUMBER: 60/061,536
; EARLIER FILING DATE: 1997-10-09
; EARLIER APPLICATION NUMBER: 60/061,532
; EARLIER FILING DATE: 1997-10-09
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 205
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-288-143-205

Query Match      58.7% Score 44: DB 4: Length 75:
Best Local Similarity 66.7% Pred. No. 3.7;
Matches      8; Conservative      3; Mismatches      1; Indels      0; Gaps      0;

OY      2 SSILCSPPLSLSLV 13
DB      15 TSQLCSPMPSLN 26

RESULT 2
US-09-177-249-208
; Sequence 208, Application US/09177249
; Patent No. 6229064
; GENERAL INFORMATION:
; APPLICANT: Fischer, Robert L.
; APPLICANT: Ohad, Nir
; APPLICANT: Kiyosue, Tomohiro
; APPLICANT: Vadegari, Ramin
```

APPLICANT: Margossian, Linda  
APPLICANT: Hairada, John  
APPLICANT: Goldberg, Robert B.  
APPLICANT: The Regents of the University of California  
TITLE OF INVENTION: Nucleic Acids That Control Seed and Fruit  
TITLE OF INVENTION: Development In Plants  
FILE REFERENCE: 023070-0861205  
CURRENT APPLICATION NUMBER: US/09/177,249  
CURRENT FILING DATE: 1998-10-22  
EARLIER APPLICATION NUMBER: US 09/071,838  
EARLIER FILING DATE: 1998-05-01  
NUMBER OF SEQ ID NOS: 324  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 208  
LENGTH: 106  
TYPE: PRT  
ORGANISM: Arabidopsis sp.  
US-09-177-249-208

Query Match 57.3%; Score 43; DB 4; Length 106;  
Best Local Similarity 72.7%; Pred. No. 7.7;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 ILCSPLPSLSL 14  
:|||||  
Db 61 MLCSPSPSLHL 71

RESULT 3  
US-09-022-983-2  
Sequence 2, Application US/09022983  
Patent No. 6159731

GENERAL INFORMATION:  
APPLICANT: Yang, Xiaolu  
APPLICANT: Khosravi-Far, Roya  
APPLICANT: Chang, Howard Y.  
APPLICANT: Baltimore, David  
TITLE OF INVENTION: DAXX, A NOVEL FAS-BINDING  
TITLE OF INVENTION: PROTEIN THAT ACTIVATES JNK AND APOPTOSIS  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
STREET: 600 Atlantic Avenue  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02210

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/022,983  
FILING DATE:

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/037,919  
FILING DATE: 12-FEB-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/051,753  
FILING DATE: 26-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Van Amsterdam, John R.  
REGISTRATION NUMBER: 40,212  
REFERENCE/DOCKET NUMBER: M0656/7036  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-720-3500  
TELEFAX: 617-720-2441

INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 739 amino acids  
TYPE: amino acid

STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
US-09-022-983-2

Query Match 56.0%; Score 42; DB 4; Length 739;  
Best Local Similarity 69.2%; Pred. No. 90;  
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 SSILCSPLPSLSL 14  
:|||||  
Db 694 TSLCSPLPSLHL 706

RESULT 4  
US-09-328-869-2  
Sequence 2, Application US/09328869B  
Patent No. 6168933

GENERAL INFORMATION:  
APPLICANT: Kaser, Matthew, R.  
APPLICANT: Hillman, Jennifer, L.  
APPLICANT: Baughn, Mariah, R.  
TITLE OF INVENTION: PHOSPHOLIPID TRANSFER PROTEIN  
FILE REFERENCE: PC-0003 US  
CURRENT APPLICATION NUMBER: US/09/328,869B  
CURRENT FILING DATE: 1999-06-08  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: PERL Program  
SEQ ID NO 2  
LENGTH: 291  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE: -  
OTHER INFORMATION: 1772859CD1  
US-09-328-869-2

Query Match 52.0%; Score 39; DB 4; Length 291;  
Best Local Similarity 88.9%; Pred. No. 95;  
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 SPLPSLSLS 15  
:|||||  
Db 242 SPLPSLSLS 250

RESULT 5  
US-09-629-774A-2  
Sequence 2, Application US/09629774A  
Patent No. 6287786

GENERAL INFORMATION:  
APPLICANT: Kaser, Matthew R.  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Baughn, Mariah R.  
TITLE OF INVENTION: PHOSPHOLIPID TRANSFER PROTEIN  
FILE REFERENCE: PC-0003-1 CIP  
CURRENT APPLICATION NUMBER: US/09/629,774A  
CURRENT FILING DATE: 2000-07-31  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: PERL Program  
SEQ ID NO 2  
LENGTH: 291  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc-feature  
OTHER INFORMATION: Incyte ID No. 6287786 1772859CD1  
US-09-629-774A-2

Query Match 52.0%; Score 39; DB 4; Length 291;  
Best Local Similarity 88.9%; Pred. No. 95;  
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 7 SPLPSLS 15  
DB 242 SPLPSLS 250

## RESULT 6

US-08-299-187-13  
Sequence 13, Application US/08299187

Patent No. 5736325

## GENERAL INFORMATION:

APPLICANT: Manowitz, Paul  
APPLICANT: Poretz, Ronald D.

APPLICANT: Park, David

APPLICANT: Rickelits, Michael H.

TITLE OF INVENTION: MARKER FOR INDIVIDUALS SUSCEPTIBLE TO

TITLE OF INVENTION: ALCOHOLISM

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESSEE: Klauber & Jackson

STREET: 411 Hackensack Avenue

CITY: Hackensack

STATE: New Jersey

COUNTRY: USA

ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/299,187

FILING DATE: 31-AUG-1994

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.

REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET NUMBER: 601-1-028

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201 487-5800

TELEFAX: 201 343-1684

TELEX: 133521

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 13 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: Internal

ORGANISM: Homo sapiens

INDIVIDUAL ISOLATE: peptide from mutant arylsulphatase A

US-08-299-187-13

Query Match 50.7%; Score 38; DB 1; Length 13;

Best Local Similarity 66.7%; Pred. No. 4.6;

Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 6 CSPSPSLSL 14  
DB 1 CAPLPSVTL 9

## RESULT 7

PCT-US95-11114-13

Sequence 13, Application PC/TUS9511114

GENERAL INFORMATION:

APPLICANT: Manowitz, Paul

APPLICANT: Poretz, Ronald D.

APPLICANT: Park, David

APPLICANT: Rickelits, Michael H.

TITLE OF INVENTION: MARKER FOR INDIVIDUALS SUSCEPTIBLE TO  
TITLE OF INVENTION: ALCOHOLISM  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Klauber & Jackson

STREET: 411 Hackensack Avenue

CITY: Hackensack

STATE: New Jersey

COUNTRY: USA

ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/11114

FILING DATE: 30-AUG-1995

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/

FILING DATE: 21-JUN-1995

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/299,187

FILING DATE: 31-AUG-1994

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.

REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET NUMBER: 1158-1-001PCT

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201 487-5800

TELEFAX: 201 343-1684

TELEX: 133521

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 13 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: Internal

ORGANISM: Homo sapiens

INDIVIDUAL ISOLATE: peptide from mutant arylsulphatase A

PCT-US95-11114-13

Query Match 50.7%; Score 38; DB 5; Length 13;

Best Local Similarity 66.7%; Pred. No. 4.6;

Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 6 CSPSPSLSL 14  
DB 1 CAPLPSVTL 9

## RESULT 8

US-07-956-700B-84

Sequence 84, Application US/07956700B

Patent No. 5539092

GENERAL INFORMATION:

APPLICANT: Robert Haselkorn and Piotr Gornicki

TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA

TITLE OF INVENTION: Carboxylase

NUMBER OF SEQUENCES: 116

CORRESPONDENCE ADDRESS:

ADDRESSEE: Arnold, White & Durkee

STREET: 321 No. 5539092th Clark Street

CITY: Chicago

STATE: Illinois

COUNTRY: USA  
ZIP: 60610  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII-DOS  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/956,700B  
FILING DATE: 19921002  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Thomas E. No. 539092thrup  
REGISTRATION NUMBER: 33,268  
REFERENCE/DOCKET NUMBER: ARCD:058  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 1-312-744-0090  
TELEFAX: 1-312-755-4489  
INFORMATION FOR SEQ ID NO: 84:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 251 amino acids  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
MOLECULE TYPE: Peptide  
US-07-956-700B-84

Query Match 50.7%; Score 38; DB 1; Length 251;  
Best Local Similarity 43.8%; Pred. No. 1.1e+02;  
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 SSSIICPLPSLSV 16  
DB 181 TSSVLRSPMPGVAV 196

RESULT 9  
US-08-476-537-84  
Sequence 84, Application US/08476537.  
Patent No. 5756290  
GENERAL INFORMATION:  
APPLICANT: Robert Haselkorn and Piotr Gornicki  
TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA  
NUMBER OF SEQUENCES: 116  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: 321 No. 5756290th Clark Street  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60610  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII-DOS  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/476,537  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/956,700  
FILING DATE: 10/21/92  
ATTORNEY/AGENT INFORMATION:  
NAME: Thomas E. No. 5756290thrup  
REGISTRATION NUMBER: 33,268  
REFERENCE/DOCKET NUMBER: ARCD:058  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 1-312-744-0090  
TELEFAX: 1-312-755-4489  
INFORMATION FOR SEQ ID NO: 84:  
SEQUENCE CHARACTERISTICS:

LENGTH: 251 amino acids  
TYPE: Amino acid  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
MOLECULE TYPE: Peptide  
US-08-476-537-84

Query Match 50.7%; Score 38; DB 1; Length 251;  
Best Local Similarity 43.8%; Pred. No. 1.1e+02;  
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 SSSIICPLPSLSV 16  
DB 181 TSSVLRSPMPGVAV 196

RESULT 10  
US-08-485-607-84  
Sequence 84, Application US/08485607  
Patent No. 5792627  
GENERAL INFORMATION:  
APPLICANT: Robert Haselkorn and Piotr Gornicki  
TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA  
NUMBER OF SEQUENCES: 116  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: 321 No. 5792627th Clark Street  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60610  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII-DOS  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/485,607  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/956,700  
FILING DATE: 10/21/92  
ATTORNEY/AGENT INFORMATION:  
NAME: Thomas E. No. 5792627thrup  
REGISTRATION NUMBER: 33,268  
REFERENCE/DOCKET NUMBER: ARCD:058  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 1-312-744-0090  
TELEFAX: 1-312-755-4489  
INFORMATION FOR SEQ ID NO: 84:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 251 amino acids  
TYPE: Amino acid  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
MOLECULE TYPE: Peptide  
US-08-485-607-84

Query Match 50.7%; Score 38; DB 1; Length 251;  
Best Local Similarity 43.8%; Pred. No. 1.1e+02;  
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 SSSIICPLPSLSV 16  
DB 181 TSSVLRSPMPGVAV 196

RESULT 11  
US-08-475-879-84  
Sequence 84, Application US/08475879.  
Patent No. 5972644

Patent No. 5972644 5786170  
GENERAL INFORMATION:  
APPLICANT: Robert Haselkorn and Piotr Gornicki  
TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA  
NUMBER OF INVENTION: Carboxylase  
NUMBER OF SEQUENCES: 116  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: 321 No. 5972644 5786170th Clark Street  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60610  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII-DOS  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/475,879  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/956,700  
FILING DATE: 10/21/92  
ATTORNEY/AGENT INFORMATION:  
NAME: Thomas E. No. 5972644 5786170chrup  
REGISTRATION NUMBER: 33,268  
REFERENCE/DOCKET NUMBER: ARCD:058  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 1-312-744-0090  
TELEFAX: 1-312-755-4489  
INFORMATION FOR SEQ. ID NO: 84:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 251 amino acids  
TYPE: Amino acid  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
MOLECULE TYPE: Peptide  
US-08-475-879-84

Query Match 50.7%; Score 38; DB 2; Length 251;  
Best Local Similarity 43.8%; Pred. No. 1.1e+02;  
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 SSSIICSPPLSLSLV 16  
DB 181 TSSVLKSPMPGVVAV 196

RESULT 12  
US-09-433-043B-84  
Sequence 84, Application US/09433043B  
Patent No. 6399342  
GENERAL INFORMATION:  
APPLICANT: HASELKORN, ROBERT  
APPLICANT: GORNICKI, PIOTR  
TITLE OF INVENTION: CYANOBACTERIAL AND PLANT ACETYL-COA CARBOXYLASE  
FILE REFERENCE: ARCD:338US  
CURRENT APPLICATION NUMBER: US/09/433,043B  
CURRENT FILING DATE: 1999-10-25  
PRIOR APPLICATION NUMBER: 08/475,879  
PRIOR FILING DATE: 1995-06-07  
PRIOR APPLICATION NUMBER: 07/956,700  
PRIOR FILING DATE: 1992-10-02  
NUMBER OF SEQ ID NOS: 128  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 84  
LENGTH: 251  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic

OTHER INFORMATION: Peptide  
US-09-433-043B-84

Query Match 50.7%; Score 38; DB 4; Length 251;  
Best Local Similarity 43.8%; Pred. No. 1.1e+02;  
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 SSSIICSPPLSLSLV 16  
DB 181 TSSVLKSPMPGVVAV 196

RESULT 13  
US-08-476-008-50  
Sequence 50, Application US/08476008  
Patent No. 5627061  
GENERAL INFORMATION:  
APPLICANT: Barry, Gerard F.  
APPLICANT: Kishore, Ganesh M.  
APPLICANT: Padgett, Stephen R.  
APPLICANT: Stallings, William C.  
TITLE OF INVENTION: Glyphosate Tolerant  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BBAF  
STREET: 700 Chesterfield Village Parkway  
CITY: St. Louis  
STATE: Missouri  
COUNTRY: USA  
ZIP: 63198  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/476,008  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/306,063  
FILING DATE: 13-SEP-1994  
APPLICATION NUMBER: US 07/749,611  
FILING DATE: 28-AUG-1991  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/576,537  
FILING DATE: 31-AUG-1990  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Hoerner Jr., Dennis R.  
REGISTRATION NUMBER: 30,914  
REFERENCE/DOCKET NUMBER: 38-21(10660)A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314)537-6099  
TELEFAX: (314)537-6047  
INFORMATION FOR SEQ. ID NO: 50:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 460 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-476-008-50

Query Match 50.7%; Score 38; DB 1; Length 460;  
Best Local Similarity 46.7%; Pred. No. 2.2e+02;  
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 SSSIICSPPLSLSLV 15  
DB 14 SSVVICAPPSKSSIS 28

RESULT 14  
US-08-306-063-50  
; Sequence 50, Application US/08306063  
; Patent No. 5633435  
; GENERAL INFORMATION:  
; APPLICANT: Barry, Gerard F.  
; APPLICANT: Kishore, Ganesh M.  
; APPLICANT: Padgett, Stephen R.  
; APPLICANT: Stallings, William C.  
; TITLE OF INVENTION: Glyphosate Tolerant  
; TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases  
; NUMBER OF SEQUENCES: 69  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BBAF  
; STREET: 700 Chesterfield Village Parkway  
; City: St. Louis  
; STATE: Missouri  
; COUNTRY: USA  
; ZIP: 63198  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/306,063  
; FILING DATE: 13-SEP-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/749,611  
; FILING DATE: 28-AUG-1991  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 31-AUG-1990  
; FILING DATE: 31-AUG-1990  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hoerner Jr., Dennis R.  
; REGISTRATION NUMBER: 30,914  
; REFERENCE/DOCKET NUMBER: 38-21(10660)A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (314)537-6099  
; TELEFAX: (314)537-6047  
; INFORMATION FOR SEQ ID NO: 50:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 460 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-306-063-50

Query Match 50.7%; Score 38; DB 1; Length 460;  
Best Local Similarity 46.7%; Pred. No. 2.2e+02;  
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 SSSIICPLPISLS 15  
||::||:| | | | |  
Db 14 SSVNICAPPSKSSIS 28

RESULT 15  
US-08-833-485-50  
; Sequence 50, Application US/08833485  
; Patent No. 5804425  
; GENERAL INFORMATION:  
; APPLICANT: Barry, Gerard F.  
; APPLICANT: Kishore, Ganesh M.  
; APPLICANT: Padgett, Stephen R.  
; APPLICANT: Stallings, William C.  
; TITLE OF INVENTION: Glyphosate Tolerant  
; TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases  
; NUMBER OF SEQUENCES: 69

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BBAF  
STREET: 700 Chesterfield Village Parkway  
City: St. Louis  
STATE: Missouri  
COUNTRY: USA  
ZIP: 63198  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/833,485  
FILING DATE: 07-APR-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/306,063  
FILING DATE: 13-SEP-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/749,611  
FILING DATE: 28-AUG-1991  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/576,537  
FILING DATE: 31-AUG-1990  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Hoerner Jr., Dennis R.  
REGISTRATION NUMBER: 30,914  
REFERENCE/DOCKET NUMBER: 38-21(15117)A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314)737-6099  
TELEFAX: (314)737-6047  
INFORMATION FOR SEQ ID NO: 50:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 460 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-833-485-50

Query Match 50.7%; Score 38; DB 1; Length 460;  
Best Local Similarity 46.7%; Pred. No. 2.2e+02;  
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 SSSIICPLPISLS 15  
||::||:| | | | |  
Db 14 SSVNICAPPSKSSIS 28

Search completed: June 25, 2003, 17:18:07  
Job time: 1.85906 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 25, 2003, 17:15:40 ; Search time 1.32438 Seconds  
(without alignments)  
1307.255 Million cell updates/sec

Title: US-10-075-846-4\_COPY\_192\_207

Sequence: 1 SSSILCSPPLSLSLV 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 417779 seqs, 108206813 residues

Total number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published\_Applications\_AA.\*  
1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	75	100.0	16	9 US-10-075-846-74	Sequence 74, Appl
2	75	100.0	312	9 US-10-075-846-15	Sequence 15, Appl
3	75	100.0	431	9 US-10-075-846-4	Sequence 4, Appl
4	47	62.7	83	9 US-09-764-891-3412	Sequence 3412, Ap
5	44	58.7	75	9 US-10-150-111-205	Sequence 205, App
6	44	58.7	508	10 US-09-771-161A-102	Sequence 102, App
7	44	58.7	756	10 US-09-771-161A-101	Sequence 101, App
8	44	58.7	887	10 US-09-771-161A-192	Sequence 192, App
9	44	57.3	887	10 US-09-771-161A-193	Sequence 193, App
10	43	57.3	106	9 US-10-213-512-208	Sequence 208, App
11	43	57.3	106	9 US-09-071-838-208	Sequence 208, App
12	43	57.3	3594	9 US-10-150-821-4	Sequence 4, Appl
13	43	57.3	3594	9 US-09-911-842-4	Sequence 4, Appl
14	41	54.7	216	10 US-09-747-155-49	Sequence 49, Appl
15	41	54.7	310	9 US-10-001-857-196	Sequence 196, App
16	40	53.3	141	10 US-09-746-284-23	Sequence 23, Appl
17	40	53.3	265	10 US-09-764-864-1006	Sequence 1006, Ap
18	40	53.3	400	10 US-09-966-871-86	Sequence 86, Appl
19	40	53.3	400	12 US-10-039-645-86	Sequence 86, Appl

20	39	52.0	68	9 US-10-106-698-5490	Sequence 5490, Ap
21	39	52.0	123	9 US-09-764-891-2743	Sequence 2743, Ap
22	39	52.0	178	10 US-09-811-284-256	Sequence 256, App
23	39	52.0	218	9 US-09-764-891-3395	Sequence 3395, Ap
24	39	52.0	270	10 US-09-841-132-404	Sequence 404, App
25	39	52.0	347	10 US-09-925-301-985	Sequence 985, App
26	39	52.0	351	9 US-10-106-698-4697	Sequence 4697, Ap
27	39	52.0	359	9 US-10-205-823-60	Sequence 60, App
28	38	50.7	25	10 US-09-864-761-36023	Sequence 36023, A
29	38	50.7	41	10 US-09-864-761-47002	Sequence 47002, A
30	38	50.7	76	9 US-10-106-698-5785	Sequence 5785, Ap
31	38	50.7	105	10 US-09-739-907-157	Sequence 157, App
32	38	50.7	141	10 US-09-864-761-37903	Sequence 37903, A
33	38	50.7	346	9 US-10-085-108-22	Sequence 22, App
34	38	50.7	362	9 US-09-796-753-142	Sequence 142, App
35	38	50.7	460	9 US-09-464-099A-50	Sequence 50, Appl
36	38	50.7	460	10 US-09-861-666-50	Sequence 50, Appl
37	38	50.7	491	9 US-09-965-529-12	Sequence 12, Appl
38	38	50.7	621	10 US-09-733-300-4	Sequence 4, Appl
39	38	50.7	667	10 US-09-816-127-1	Sequence 1, Appl
40	38	50.7	827	9 US-10-101-464A-915	Sequence 915, App
41	38	50.7	897	9 US-10-138-927-45	Sequence 45, Appl
42	38	50.7	943	9 US-09-933-261-7	Sequence 7, Appl
43	38	50.7	943	9 US-10-256-702-7	Sequence 7, Appl
44	37.5	50.0	126	10 US-09-740-668A-42	Sequence 42, Appl
45	37.5	50.0	300	9 US-10-174-590-548	Sequence 548, App

#### ALIGNMENTS

RESULT 1  
US-10-075-846-74  
; Sequence 74, Application US/10075846  
; Publication No. US2003032608A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING A NOVEL GLYCINE RECEPTOR ALPHA SUBUNIT  
; FILE REFERENCE: D0079 NP  
; CURRENT APPLICATION NUMBER: US/10/075, 846  
; PRIOR FILING DATE: 2002-02-13  
; PRIOR APPLICATION NUMBER: US 60/269, 535  
; NUMBER OF SEQ ID NOS: 81  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 74  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-075-846-74

Query Match 100.0%; Score 75; DB 9; Length 16;  
Best Local Similarity 100.0%; Pred. No. 3.1e-05;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 SSSILCSPPLSLSLV 16  
DB 1 SSSILCSPPLSLSLV 16

RESULT 2  
US-10-075-846-15  
; Sequence 15, Application US/10075846  
; Publication No. US2003032608A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING A NOVEL GLYCINE RECEPTOR ALPHA SUBUNIT  
; FILE REFERENCE: D0079 NP  
; CURRENT APPLICATION NUMBER: US/10/075, 846  
; PRIOR FILING DATE: 2002-02-13  
; PRIOR APPLICATION NUMBER: US 60/269, 535

;; PRIOR FILING DATE: 2001-02-16  
;; NUMBER OF SEQ ID NOS: 81  
;; SOFTWARE: PatentIn version 3.0  
;; SEQ ID NO 15  
;; LENGTH: 312  
;; TYPE: PRT  
;; ORGANISM: homo sapiens  
US-10-075-846-15

Query Match 100.0%; Score 75; DB 9; Length 312;  
Best Local Similarity 100.0%; Pred. No. 0.00061;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSSIICSPPLSLSLV 16  
Db 149 SSSIICSPPLSLSLV 164

RESULT 3  
US-10-075-846-4  
; Sequence 4, Application US/10075846  
; Publication No. US20030032608A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING A NOVEL GLYCINE RECEPTOR ALPHA SUBUNIT B  
; FILE REFERENCE: D0079 NP  
; CURRENT APPLICATION NUMBER: US/10/075,846  
; CURRENT FILING DATE: 2002-02-13  
; PRIOR APPLICATION NUMBER: US 60/269,535  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 81  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 4  
; LENGTH: 431  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-10-075-846-4

Query Match 100.0%; Score 75; DB 9; Length 431;  
Best Local Similarity 100.0%; Pred. No. 0.00084;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSSIICSPPLSLSLV 16  
Db 192 SSSIICSPPLSLSLV 207

RESULT 4  
US-09-764-891-3412  
; Sequence 3412, Application US/09764891  
; Publication No. US20030077808A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC006  
; CURRENT APPLICATION NUMBER: US/09/764,891  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 10231  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3412  
; LENGTH: 83  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (51)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-764-891-3412

Query Match 62.7%; Score 47; DB 9; Length 83;  
Best Local Similarity 73.3%; Pred. No. 3;

Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 1 SSSIICSPPLSLSLV 15  
Db 66 SSSCLVSPPLSLFLA 80

RESULT 5  
US-10-150-111-205  
; Sequence 205, Application US/10150111  
; Publication No. US20030078386A1  
; GENERAL INFORMATION:  
; APPLICANT: Rubin et al.  
; TITLE OF INVENTION: Secreted Protein HPEAD48  
; FILE REFERENCE: P2018P1D1  
; CURRENT APPLICATION NUMBER: US/10/150,111  
; CURRENT FILING DATE: 2002-05-20  
; PRIOR APPLICATION NUMBER: 09/288,143  
; PRIOR FILING DATE: 1999-04-08  
; PRIOR APPLICATION NUMBER: PCT/US98/21142  
; PRIOR FILING DATE: 1998-10-08  
; PRIOR APPLICATION NUMBER: 60/061,463  
; PRIOR FILING DATE: 1997-10-09  
; PRIOR APPLICATION NUMBER: 60/061,529  
; PRIOR FILING DATE: 1997-10-09  
; PRIOR APPLICATION NUMBER: 60/071,498  
; PRIOR FILING DATE: 1997-10-09  
; PRIOR APPLICATION NUMBER: 60/061,527  
; PRIOR FILING DATE: 1997-10-09  
; PRIOR APPLICATION NUMBER: 60/061,536  
; PRIOR FILING DATE: 1997-10-09  
; PRIOR APPLICATION NUMBER: 60/061,532  
; PRIOR FILING DATE: 1997-10-09  
; NUMBER OF SEQ ID NOS: 219  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 205  
; LENGTH: 75  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-150-111-205

Query Match 58.7%; Score 44; DB 9; Length 75;  
Best Local Similarity 66.7%; Pred. No. 7.9;  
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 SSSICSPPLSLSLV 13  
Db 15 TSQCLSPMPSLN 26

RESULT 6  
US-09-771-161A-102  
; Sequence 102, Application US/09771161A  
; Patent No. US2002010811A1  
; GENERAL INFORMATION:  
; APPLICANT: Levine, et al.  
; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES  
; FILE REFERENCE: 802620-2005.1  
; CURRENT APPLICATION NUMBER: US/09/771,161A  
; CURRENT FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: 09/724,676  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: 136776  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: 135619  
; PRIOR FILING DATE: 2000-04-12  
; NUMBER OF SEQ ID NOS: 273  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 102  
; LENGTH: 508  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-771-161A-102

Query Match 58.7%; Score 44; DB 10; Length 508;  
Best Local Similarity 69.2%; Pred. No. 54;  
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SSSICSPPLSPS 13  
111:11111:1  
DB 442 SSOITSPPLSPVS 454

## RESULT 7

US-09-771-161A-101  
; Sequence 101, Application US/09771161A  
; Patent No. US20020110811A1  
; GENERAL INFORMATION:  
; APPLICANT: LEVINE, et al.  
; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES  
; FILE REFERENCE: 802620-2005.1  
; CURRENT APPLICATION NUMBER: US/09/771.161A  
; PRIOR FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: 09/724,676  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: 136776  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: 135619  
; PRIOR FILING DATE: 2000-04-12  
; NUMBER OF SEQ ID NOS: 273  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 101  
; LENGTH: 756  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-771-161A-101

Query Match 58.7%; Score 44; DB 10; Length 756;  
Best Local Similarity 69.2%; Pred. No. 80;

Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SSSICSPPLSPS 13  
111:11111:1  
DB 442 SSOITSPPLSPVS 454

## RESULT 8

US-09-771-161A-192  
; Sequence 192, Application US/09771161A  
; Patent No. US20020110811A1  
; GENERAL INFORMATION:  
; APPLICANT: LEVINE, et al.  
; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES  
; FILE REFERENCE: 802620-2005.1  
; CURRENT APPLICATION NUMBER: US/09/771.161A  
; PRIOR FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: 09/724,676  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: 136776  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: 135619  
; PRIOR FILING DATE: 2000-04-12  
; NUMBER OF SEQ ID NOS: 273  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 192  
; LENGTH: 887  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-771-161A-192

Query Match 58.7%; Score 44; DB 10; Length 887;  
Best Local Similarity 69.2%; Pred. No. 94;

Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SSSICSPPLSPS 13  
111:11111:1

DB 442 SSOITSPPLSPVS 454

## RESULT 9

US-09-771-161A-193  
; Sequence 193, Application US/09771161A  
; Patent No. US20020110811A1  
; GENERAL INFORMATION:  
; APPLICANT: LEVINE, et al.  
; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES  
; FILE REFERENCE: 802620-2005.1  
; CURRENT APPLICATION NUMBER: US/09/771.161A  
; PRIOR FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: 09/724,676  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: 136776  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: 135619  
; PRIOR FILING DATE: 2000-04-12  
; NUMBER OF SEQ ID NOS: 273  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 193  
; LENGTH: 887  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-771-161A-193

Query Match 58.7%; Score 44; DB 10; Length 887;  
Best Local Similarity 69.2%; Pred. No. 94;

Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SSSICSPPLSPS 13  
111:11111:1  
DB 442 SSOITSPPLSPVS 454

## RESULT 10

US-10-213-512-208  
; Sequence 208, Application US/10213512  
; Publication No. US20030110536A1  
; GENERAL INFORMATION:  
; APPLICANT: Fischer, Robert L.  
; APPLICANT: Ohad, Nir  
; APPLICANT: Kiyosue, Tomohiro  
; APPLICANT: Yadegari, Ramon  
; APPLICANT: Margossian, Linda  
; APPLICANT: Harada, John  
; APPLICANT: Goldberg, Robert B.  
; TITLE OF INVENTION: The Regents of the University of California  
; TITLE OF INVENTION: Combinations of Nucleic Acids That Control Seed and  
; FILE REFERENCE: 023070-086100S  
; CURRENT APPLICATION NUMBER: US/10/213.512  
; PRIOR FILING DATE: 2002-08-06  
; PRIOR APPLICATION NUMBER: US/09/177.206  
; PRIOR FILING DATE: 1998-10-22  
; PRIOR APPLICATION NUMBER: US 09/071.838  
; PRIOR FILING DATE: 1998-05-01  
; NUMBER OF SEQ ID NOS: 324  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 208  
; LENGTH: 106  
; TYPE: PRT  
; ORGANISM: Arabidopsis sp.  
US-10-213-512-208

Query Match 57.3%; Score 43; DB 9; Length 106;  
Best Local Similarity 72.7%; Pred. No. 16;

Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 ILCSPLPSLSL 14  
:1111111  
DB 61 MLCSPPLSLHL 71

RESULT 11  
US-09-071-838-208  
Sequence 208, Application US/09071838  
Patent No. US20020152501A1  
GENERAL INFORMATION:  
APPLICANT: Fischer, Robert L.  
APPLICANT: Ohad, Nir  
APPLICANT: Kiyosue, Tomohiro  
APPLICANT: Yadegari, Ramlin  
APPLICANT: Margossian, Linda  
APPLICANT: Goldberg, Robert B.  
TITLE OF INVENTION: Nucleic Acids That Control Seed and  
TITLE OF INVENTION: Fruit Development in Plants  
NUMBER OF SEQUENCES: 324  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/071,838  
FILING DATE: 01-MAY-1998  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 023070-08610005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 208:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 106 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-071-838-208

Query Match 57.3%; Score 43; DB 10; Length 106;  
Best Local Similarity 72.7%; Pred. No. 16;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 IICSPPLSL 14  
DB 61 MLCSPPLSL 71

RESULT 12  
US-10-150-821-4  
Sequence 4, Application US/10150821  
Patent No. US20020192758A1  
GENERAL INFORMATION:  
APPLICANT: Weicher, Andrew A.  
APPLICANT: Elliott, Gary S.  
TITLE OF INVENTION: C3B/C4B COMPLEMENT RECEPTOR-LIKE MOLECULES AND USES  
TITLE OF INVENTION: THEREOF  
FILE REFERENCE: 01017/37592  
CURRENT APPLICATION NUMBER: US/10/150,821  
CURRENT FILING DATE: 2002-05-16  
PRIOR APPLICATION NUMBER: US/09/911,842  
PRIOR FILING DATE: 2001-07-24  
PRIOR APPLICATION NUMBER: US 60/222,438  
PRIOR FILING DATE: 2000-08-01

NUMBER OF SEQ ID NOS: 5  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 4  
LENGTH: 3594  
TYPE: PRT  
ORGANISM: Mus musculus  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1757)..()  
OTHER INFORMATION: Xaa = any or unknown amino acid  
US-10-150-821-4

Query Match 57.3%; Score 43; DB 9; Length 3594;  
Best Local Similarity 62.5%; Pred. No. 5.4e+02;  
Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 SSIICSPPLSLSV 16  
DB 1649 SSMFCDCPSLEGSV 1664

RESULT 13  
US-09-911-842-4  
Sequence 4, Application US/09911842  
Patent No. US20020151483A1  
GENERAL INFORMATION:  
APPLICANT: Weicher, Andrew A.  
APPLICANT: Elliott, Gary S.  
TITLE OF INVENTION: C3B/C4B COMPLEMENT RECEPTOR-LIKE MOLECULES AND USES THEREOF  
FILE REFERENCE: 01017/37592  
CURRENT APPLICATION NUMBER: US/09/911,842  
CURRENT FILING DATE: 2001-07-24  
PRIOR APPLICATION NUMBER: US 60/222,438  
PRIOR FILING DATE: 2000-08-01  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 4  
LENGTH: 3594  
TYPE: PRT  
ORGANISM: Mus musculus  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1757)..()  
OTHER INFORMATION: Xaa = any or unknown amino acid  
US-09-911-842-4

Query Match 57.3%; Score 43; DB 10; Length 3594;  
Best Local Similarity 62.5%; Pred. No. 5.4e+02;  
Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 SSIICSPPLSLSV 16  
DB 1649 SSMFCDCPSLEGSV 1664

RESULT 14  
US-09-747-155-49  
Sequence 49, Application US/09747155  
Patent No. US20020151692A1  
GENERAL INFORMATION:  
APPLICANT: Rouquier, Sylvie  
APPLICANT: Giorgi, Dominique  
TITLE OF INVENTION: No. US20020151692A1el Polypeptides and Nucleic Acids Encoding  
FILE REFERENCE: 19904-008 (C009B6834US)  
CURRENT APPLICATION NUMBER: US/09/747,155  
CURRENT FILING DATE: 2000-12-21  
PRIOR APPLICATION NUMBER: 60/171,746  
PRIOR FILING DATE: 1999-12-22  
NUMBER OF SEQ ID NOS: 431  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 49  
LENGTH: 216  
TYPE: PRT

Page 5

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 25, 2003, 17:13:55 ; Search time 1.07383 Seconds  
(without alignments)  
1432.403 Million cell updates/sec

Title: US-10-075-846-4\_COPY\_192\_207

Perfect score: 75  
Sequence: 1 SSSIICSPPLPSLSLV 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	65.3	490	2 T49096	hypothetical prote
2	46	61.3	117	2 C85683	probable IS encode
3	44	58.7	887	1 S57219	1-phosphatidylinos
4	42	56.0	91	1 MNVUAV	nonstructural prot
5	42	56.0	114	2 A72535	hypothetical prote
6	42	56.0	141	2 B82845	hypothetical prote
7	42	56.0	478	2 H71345	probable polynucle
8	42	56.0	1010	2 T41077	hypothetical struc
9	41	54.7	396	2 T02483	probable protein p
10	41	54.7	944	2 T38130	probable helicase
11	41	54.7	2123	2 F86348	hypothetical prote
12	40	53.3	247	2 T22881	hypothetical prote
13	40	53.3	262	2 D71557	hypothetical prote
14	40	53.3	313	2 D96616	hypothetical prote
15	40	53.3	798	2 S09867	hypothetical prote
16	40	53.3	965	2 T38430	hypothetical prote
17	39.5	52.7	227	2 S35240	hypothetical prote
18	39	52.0	261	2 E81520	heat shock protein
19	39	52.0	270	2 C72019	conserved hypotet
20	39	52.0	270	2 B86604	CT764 hypothetical
21	39	52.0	449	2 T16259	hypothetical prote
22	39	52.0	509	2 A36392	segmentation prote
23	39	52.0	647	2 A37086	beta-galactosidase
24	39	52.0	740	2 S55140	hypothetical prote
25	39	52.0	841	1 S24462	probable 3',5'-cyc
26	39	52.0	918	1 D88544	protein R08D7 6 [1
27	38	50.7	100	2 AG0379	acetylactate synth
28	38	50.7	107	2 A72456	hypothetical prote
29	38	50.7	120	2 AC2286	NADH dehydrogenase

30	38	50.7	124	2 A54773	Ac1 protein - mous
31	38	50.7	133	2 A85093	hypothetical prote
32	38	50.7	141	2 F71555	hypothetical prote
33	38	50.7	185	2 T10677	pathogenesis-relat
34	38	50.7	201	2 G68736	hypothetical prote
35	38	50.7	231	2 A82538	hypothetical prote
36	38	50.7	297	2 T27206	hypothetical prote
37	38	50.7	306	2 T49541	hypothetical prote
38	38	50.7	328	2 JN0882	gonadotropin-relea
39	38	50.7	353	2 T36089	probable lacI-fam1
40	38	50.7	449	2 B85069	hypothetical prote
41	38	50.7	491	2 T46915	hypothetical prote
42	38	50.7	554	2 T41612	amino acid permeas
43	38	50.7	576	2 T11046	NADH2 dehydrogenas
44	38	50.7	621	2 S27752	anthranilate synth
45	38	50.7	621	2 J01685	anthranilate synth

#### ALIGNMENTS

RESULT 1  
T49096  
hypothetical protein F4F15.280 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 02-Jun-2000  
C:Accession: T49096  
R:Alcaraz, J.P.; Clabault, G.; Cottet, A.; Maché, R.; Mewes, H.W.; Rudd, S.; Lemcke, submitted to the Protein Sequence Database, April 2000  
A:Reference number: Z25015  
A:Accession: T49096  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-490 <ALC>  
A:Cross-references: EMBL:AL049711; GSPDB:GN00061; ATSP:F4F15.280  
A:Experimental source: cultivar Columbia; BAC clone F4F15  
C:Genetics:  
A:Gene: ATSP:F4F15.280  
A:Map position: 3  
A:Insertions: 70/2; 433/3

Query Match 65.3%; Score 49; DB 2; Length 490;  
Best Local Similarity 73.3%; Pred. No. 2.9;  
Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Db 1 SSSIICSPPLPSLS 15  
129 SSSIIMDPVPPLSL 143

RESULT 2  
C85683  
probable IS encoded protein Z1827 [Imported] - Escherichia coli (strain O157:H7, subs  
C:Species: Escherichia coli  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
C:Accession: C85683  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May  
hiller, L.; Grothbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apoda  
Nature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A:Reference number: AB5480; MUID:21074935; PMID:11206551  
A:Accession: C85683  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-117 <STO>  
A:Cross-references: GB:AE005174; NID:G12514746; PIDN:AAG55927.1; GSPDB:GN00145; UWGP:  
C:Experimental source: strain O157:H7, substrain EDL933  
C:Genetics:  
A:Gene: Z1827

Query Match 61.3%; Score 46; DB 2; Length 117;  
Best Local Similarity 53.3%; Pred. No. 2.2;  
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 2 SSILCSPPLSLSV 16  
 :||:||||:|  
 Db 12 TSILCSPMTSLKTSI 26

## RESULT 3

S57219  
 1-phosphatidylinositol 3-kinase (EC 2.7.1.137) Vps34-type [validated] - human  
 C:Species: Homo sapiens (man)  
 C:Date: 27-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 01-Feb-2002  
 C:Accession: S57219  
 R:Volinia, S.; Dhond, R.; Vanhaesebroeck, B.; MacDougall, L.K.; Stein, R.; Zvelebil, M.  
 EMBL J. 14, 3339-3348, 1995  
 A:Title: A human phosphatidylinositol 3-kinase complex related to the yeast Vps34p-Vps15  
 A:Reference number: S57219; MUID:95354652; PMID:7628435  
 A:Accession: S57219  
 A:Molecule type: mRNA  
 A:Residues: 1-887 <VOL>  
 A:Function:  
 A:Description: converts ATP and 1-phosphatidyl-1D-myo-inositol to ADP and 1-phosphatidyl-  
 A:Pathway: required for vacuolar sorting and segregation; involved in both internalizati  
 A>Note: specific for phosphatidylinositol, inactive on phosphatidylinositol-3-phosphate  
 C:Keywords: slime mold phosphatidylinositol 3-kinase  
 C:Keywords: phosphotransferase; protein trafficking

Query Match 58.7%; Score 44; DB 1; Length 887;  
 Best Local Similarity 69.2%; Pred. No. 34;  
 Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 SSILCSPPLSLSV 13  
 ||:||||:|  
 Db 442 SSQITSPPLSPSVS 454

## RESULT 4

MNVNAV  
 nonstructural protein NS-S - alno virus  
 C:Species: alno virus  
 C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 16-Jul-1999  
 C:Accession: S07943  
 R:Akash, H.; Gay, M.; Thara, T.; Bishop, D.H.L.  
 Virus Res. 1, 51-63, 1984  
 A:Title: Localized conserved regions of the S RNA gene products of bunyaviruses are reve  
 A:Reference number: S07414; MUID:85170502; PMID:6532000  
 A:Accession: S07943  
 A:Molecule type: genomic RNA  
 A:Residues: 1-91 <AKA>  
 A:Cross-references: EMBL:M22011; NID:9210064; PIDN:AAA42544.1; PID:9210066  
 C:Genetics:  
 A:Map position: segment 5  
 C:Superfamily: bunyavirus nonstructural protein  
 C:Keywords: nonstructural protein

Query Match 56.0%; Score 42; DB 1; Length 91;  
 Best Local Similarity 57.1%; Pred. No. 7.6;  
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 SSILCSPPLSLSV 14  
 :||:||||:|  
 Db 73 TQILCOTLPRLST 86

## RESULT 5

A72535  
 hypothetical protein APE1566 - Aeropyrum pernix (strain K1)  
 C:Species: Aeropyrum pernix  
 C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jun-2000  
 C:Accession: A72535  
 R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jin-no, K.; Takaf  
 awa, H.; Takamiya, M.; Mesuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K  
 DNA Res. 6 83-101, 1999  
 A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr

A:Reference number: A72450; MUID:99310339; PMID:10382966  
 A:Accession: A72535  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-114 <KAW>  
 A:Cross-references: DDBJ:AP000062; NID:95105244; PIDN:BAA80566.1; PID:dl044352; PID:9  
 A:Experimental source: strain K1  
 C:Genetics:  
 A:Gene: APE1566  
 C:Superfamily: Aeropyrum pernix hypothetical protein APE1566

Query Match 56.0%; Score 42; DB 2; Length 114;  
 Best Local Similarity 54.5%; Pred. No. 9.5;  
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 6 CSPPLSLSLSV 16  
 ||:||||:|  
 Db 28 CTPPSMNLST 38

## RESULT 6

B82845  
 hypothetical protein XF0131 [Imported] - Xylella fastidiosa (strain 9a5c)  
 C:Species: Xylella fastidiosa  
 C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
 C:Accession: B82845  
 R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq  
 Nature 406, 151-157, 2000  
 A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
 A:Reference number: A82515; MUID:20365717; PMID:10910347  
 A>Note: for a complete list of authors see reference number A59328 below  
 A:Accession: B82845  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-141 <STM>

A:Cross-references: GB:AE003866; GB:AE003849; NID:99104906; PIDN:AAF82944.1; GSPDB:GN  
 A:Experimental source: strain 9a5c  
 R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.  
 Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer  
 as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.  
 submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franco, S.C.; Franco, M.C.; Fr  
 J.D.; Junqueira, M.L.; Kemp, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; La  
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.L.; Marques, M.V.; Martins  
 A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.  
 F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,  
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sava  
 A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv  
 M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.  
 A:Reference number: A59328  
 A:Contents: annotation  
 C:Genetics:  
 A:Gene: XF0131

Query Match 56.0%; Score 42; DB 2; Length 141;  
 Best Local Similarity 57.1%; Pred. No. 12;  
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 SSILCSPPLSLSV 14  
 :||:||||:|  
 Db 9 TSLRCAPPLSLLL 22

## RESULT 7

H71345  
 probable polynucleotide adenylyltransferase (pcna) - syphilis spirochete  
 C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)  
 C:Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 05-Nov-1999  
 C:Accession: H71345  
 R:Riferson, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; G  
 rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; M  
 they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.  
 Science 281, 375-388, 1998



A>Title: Complete genome sequence of *Treponema pallidum*, the syphilis spirochete.  
A:Reference number: A71250; MUID:98332770; PMID:9665876  
A:Accession: H71345  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-478 <COL>  
A:Cross-references: GB:AE001208; GB:AE000520; NID:g33322538; PIDN:AAC65258.1; PID:g3332254  
A:Experimental source: strain Nichols  
C:Genetics:  
A:Gene: TP0270

Query Match 56.0%; Score 42; DB 2; Length 478;  
Best Local Similarity 53.8%; Pred. No. 39;  
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 2 SSILCSPLPSLSL 14  
DB 203 SKMLCTPRPSIAL 215

## RESULT 8

T41077

hypothetical structural protein - fission yeast (*Schizosaccharomyces pombe*)  
C:Species: Schizosaccharomyces pombe  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999

C:Accession: T41077  
R:McDougal, R.C.; Rajandream, M.A.; Barrell, B.G.; Zimmermann, W.; Wambutt, R.  
submitted to the EMBL Data Library, August 1999  
A:Reference number: Z21822

A:Accession: T41077  
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA  
A:Residues: 1-1010 <EMCD>

A:Cross-references: EMBL:AL109957; PIDN:CAB53076.1; GSPDB:GN00068; SPDB:SPCC16A11.04  
A:Experimental source: strain 972h-; cosmid c16A11  
C:Genetics:  
A:Gene: SPDB:SPCC16A11.04

A:Map position: 3  
A:Map position: 3

Query Match 56.0%; Score 42; DB 2; Length 1010;  
Best Local Similarity 62.5%; Pred. No. 81;  
Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 1 SSILCSPLPSLSL 16  
DB 572 SKINSPPSPISQISLV 587

## RESULT 9

T02483

probable protein phosphatase 2C At2g30020 [imported] - *Arabidopsis thaliana*  
N:Alternate names: hypothetical protein F23F1.6  
C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 05-Mar-1999 #sequence\_revision 05-Mar-1999 #text\_change 03-Jun-2002  
C:Accession: T02483; E84703  
R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul  
submitted to the EMBL Data Library, August 1998  
A:Description: Arabidopsis thaliana chromosome II BAC F23F1 genomic sequence.  
A:Reference number: Z14675

A:Accession: T02483  
A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA  
A:Residues: 1-396 <ROU>

A:Cross-references: EMBL:AC004680; NID:g3420043; PID:g3420049  
A:Experimental source: cultivar Columbia

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Motilal, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.;  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A>Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.  
A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: E84703  
A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-396 <STO>

A:Cross-references: GB:AE002093; NID:g3420049; PIDN:ACJ1850.1; GSPDB:GN00139  
C:Genetics:

A:Gene: At2g30020; F23F1.6  
A:Map position: 2  
A:Introns: 168/3

C:Superfamily: Arabidopsis thaliana hypothetical protein F7A7.220

Query Match 54.7%; Score 41; DB 2; Length 396;  
Best Local Similarity 66.7%; Pred. No. 47;  
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 SSILCSPLPSLSL 16  
DB 24 SSILSPQESLSLTL 38

## RESULT 10

T38130

probable helicase - fission yeast (*Schizosaccharomyces pombe*)  
C:Species: Schizosaccharomyces pombe  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999

C:Accession: T38130  
R:Badcock, K.; Churcher, C.M.; Wood, V.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, May 1997  
A:Reference number: Z21772

A:Accession: T38130  
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA  
A:Residues: 1-944 <BAD>

A:Cross-references: EMBL:Z95334; PIDN:CAB08602.1; GSPDB:GN00066; SPDB:SPAC20G8.08c  
A:Experimental source: strain 972h-; cosmid c20G8  
C:Genetics:

A:Gene: SPDB:SPAC20G8.08c  
A:Map position: 1  
A:Map position: 1

A:Introns: 538/3; 583/3; 717/3; 894/3

Query Match 54.7%; Score 41; DB 2; Length 944;  
Best Local Similarity 60.0%; Pred. No. 11e+02;  
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 1 SSILCSPLPSLSL 15  
DB 32 SSIECSPIQQLDLS 46

## RESULT 11

F86348

hypothetical protein F24J8.17 [imported] - *Arabidopsis thaliana*  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Dec-2001

C:Accession: F86348  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon

Chin, C.W.; Chung, M.K.; Con, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,  
ansen, N.F.; Hughes, B.; Hultzer, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marzla

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A>Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.  
A:Reference number: A86141; MUID:21016719; PMID:11130712  
A:Accession: F86348

A:Status: preliminary

A:Molecule type: DNA  
A:Residues: 1-2123 <STO>

A:Cross-references: GB:AE005172; NID:g9454580; PIDN:AAFG7903.1; GSPDB:GN00141  
C:Genetics:

A:Map position: 1  
Query Match 54.7%; Score 41; DB 2; Length 2123;



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 25, 2003, 17:12:59 ; Search time 0.644295 Seconds  
(without alignments)  
1029.995 Million cell updates/sec

Title: US-10-075-846-4\_COPY\_192\_207  
Perfect score: 75  
Sequence: 1 SSSILCSPLPSLSLV 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues  
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	56.0	91	1 VNST_AINOV	P12413 alno virus.
2	40	53.3	401	1 OPRM_PIG	O95247 sus scrofa
3	40	53.3	798	1 HEPB_HCMVA	P16827 human cytom
4	39.5	52.7	227	1 HS2C_ARATH	P31170 arabidopsis
5	39	52.0	509	1 RUNT_DROME	P22814 drosophila
6	39	52.0	647	1 BGAL_MOUSE	P23780 mus musculu
7	39	52.0	739	1 A4B1_MOUSE	O9wv76 mus musculu
8	39	52.0	740	1 YNO4_YEAST	P53907 saccharomyc
9	39	52.0	918	1 YNE6_CAEEL	P30645 caenorhabd1
10	38	50.7	120	1 NU3C_ANASP	O44239 anabaena sp
11	38	50.7	124	1 APC4_MOUSE	O61268 mus musculu
12	38	50.7	328	1 GRHR_SHEEP	P32237 ovis aries
13	38	50.7	464	1 IRLS_BURPS	O31396 burkholderi
14	38	50.7	554	1 YCT3_SCHPO	O59813 schizosach
15	38	50.7	621	1 TRPX_ARATH	P32069 arabidopsis
16	38	50.7	703	1 PCCA_HUMAN	P05165 homo sapien
17	38	50.7	1586	1 AROL_EMENT	P07547 e. dentafunc
18	38	50.7	4092	1 DYHC_YEAST	P36622 saccharomyc
19	37.5	50.0	220	1 RK21_ARATH	P51412 arabidopsis
20	37	49.3	168	1 PLAT_POPNI	P00299 populus nig
21	37	49.3	168	1 PLAT_POPNI	P11970 populus nig
22	37	49.3	176	1 Y069_TREPA	O83108 treponema p
23	37	49.3	288	1 VG27_HSV11	O00113 ictaluriid h
24	37	49.3	365	1 Y023_CAEEL	O09452 caenorhabd1
25	37	49.3	773	1 GLGB_STNPF	P16954 synchococc
26	37	49.3	803	1 RIR1_CRYPV	O61065 cryptospori
27	37	49.3	817	1 HUNB_MUSDO	O01778 musca domes
28	37	49.3	1229	1 KPB2_FUGRU	O9w6r1 fugu rubrip
29	37	49.3	1770	1 RIIS_YEAST	P43565 saccharomyc
30	36.5	48.7	474	1 MIG1_KLULA	P50898 klyveromyc
31	36.5	48.7	476	1 EFTU_ARATH	P17745 arabidopsis
32	36.5	48.7	1609	1 FIG2_YEAST	P23553 saccharomyc
33	36	48.0	96	1 GRO_RAT	P14095 rattus norv

34	36	48.0	108	1 SLIB_HUMAN	P01286 homo sapien
35	36	48.0	264	1 CERC_SCHMA	P12546 schistosoma
36	36	48.0	271	1 FRA1_HUMAN	P15407 homo sapien
37	36	48.0	274	1 PSBS_SPTOL	O02060 spinacia o1
38	36	48.0	275	1 FRA1_RAT	P10158 rattus norv
39	36	48.0	430	1 MS12_AGRRH	P50201 agrobacteri
40	36	48.0	440	1 FU10_ARATH	O951p6 arabidopsis
41	36	48.0	475	1 ETS6_DROME	P29776 drosophila
42	36	48.0	533	1 FUT5_ARATH	O951p4 arabidopsis
43	36	48.0	675	1 NED1_MOUSE	P33215 mus musculu
44	36	48.0	956	1 CB31_YEAST	P32504 saccharomyc
45	36	48.0	1165	1 LEPR_HUMAN	P48357 homo sapien

ALIGNMENTS

RESULT 1  
VNST\_AINOV STANDARD: PRT: 91 AA.  
AC P12413:  
DT 01-OCT-1989 (rel. 12, Created)  
DT 01-OCT-1989 (rel. 12, Last sequence update)  
DT 01-JAN-1990 (rel. 13, Last annotation update)  
DE Nonstructural protein NS-S.  
OS Alno virus.  
OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Orthobunyavirus.  
OX NCBI\_TaxID=11582;  
RN (1)  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85170502; PubMed=6532000;  
RA Akashi H., Gay M., Ihara T., Bishop D.H.L.;  
RT "Localized conserved regions of the S RNA gene products of  
RT bunyaviruses are revealed by sequence analyses of the Simbu serogroup  
RT Alno virus.";  
RL Virus Res. 1:51-63(1984).  
CC -I- SIMILARITY: 35% TO LA CROSSE BUNYAVIRUS NONSTRUCTURAL PROTEIN NS-S  
CC AND TO SNOWSHOE HARE BUNYAVIRUS NS-S.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC  
CC EMBL: M22011; AAA42544.1; -  
CC PIR: S07943; MNVUAV.  
CC InterPro: IPR000797; Bunya\_NSS.  
CC Pfam: PF01104; Bunya\_NS-S; 1.  
CC ProDom: PD002170; Bunya\_NSS; 1.  
CC Nonstructural protein.  
SQ SEQUENCE 91 AA; 10503 MW; 6B9195505744BA0 CRC64;  
  
Query Match 56.0%; Score 42; DB 1; Length 91;  
Best Local Similarity 57.1%; Pred. NO. 1.7;  
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
  
OY 1 SSSILCSPLPSLSL 14  
: ||| |||||  
Db 73 TQOILCQTLPSLSI 86  
  
RESULT 2  
OPRM\_PIG STANDARD: PRT: 401 AA.  
AC O95247:  
DT 01-NOV-1997 (rel. 35, Created)  
DT 01-NOV-1997 (rel. 35, Last sequence update)  
DT 15-JUL-1999 (rel. 38, Last annotation update)  
DE Mu-type opioid receptor (MOR-1).  
GN OPRM1.

```

OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain cortex;
RC MEDLINE=99032385; PubMed=9817447;
RA Pampusch M.S., Osinski M.A., Brown D.R., Murtaugh M.P.;
RT "The porcine mu opioid receptor: molecular cloning and mRNA
RT distribution in lymphoid tissues.";
RT J. Neuroimmunol. 90:192-198(1998).
CC -1- FUNCTION: INHIBITS NEUROTRANSMITTER RELEASE BY REDUCING CALCIUM
CC ION CURRENTS AND INCREASING POTASSIUM ION CONDUCTANCE. RECEPTOR
CC FOR BETA-ENDORPHIN.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L38645; AAB33770.1; -
CC InterPro: IPR000276; GPCR_Rhodpsn.
CC Pfam: PF00001; 7tm.1; 1.
DR PRINTS: PR00237; GPCRHOOPS.
DR PROSITE: PS00237; G_PROTEIN_RECP_FL_1; 1.
DR PROSITE: PS0262; G_PROTEIN_RECP_FL_2; 1.
KW G-protein coupled receptor; Transmembrane; glycoprotein;
KW Phosphorylation; Lipoprotein; Palmitate.
FT DOMAIN 1 67 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 68 97 1 (POTENTIAL).
FT DOMAIN 98 106 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 107 124 2 (POTENTIAL).
FT DOMAIN 125 146 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 147 166 3 (POTENTIAL).
FT DOMAIN 167 196 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 197 212 4 (POTENTIAL).
FT DOMAIN 213 237 5 (POTENTIAL).
FT TRANSMEM 238 260 5 (POTENTIAL).
FT DOMAIN 261 283 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 284 306 6 (POTENTIAL).
FT DOMAIN 307 314 7 (POTENTIAL).
FT TRANSMEM 315 331 7 (POTENTIAL).
FT DOMAIN 332 401 CYTOPLASMIC (POTENTIAL).
FT DISULFID 143 220 BY SIMILARITY.
FT LIPID 354 354 PALMITATE (POTENTIAL).
FT CARBOHYD 9 9 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 12 12 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 34 34 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 41 41 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 49 49 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 401 AA; 45098 MW; 1A208BC5ED7F83FA CRC64;

Query Match 53.3%; Score 40; DB 1; Length 401;
Best local Similarity 61.5%; Pred. No. 18;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```

```

DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE DNA helicase/primase complex associated protein.
GN UL102.
OS Human cytomegalovirus (strain AD169).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10360;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=90269039; PubMed=2161319;
RA Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,
RA Horsnell T., Hutchinson S.C., Ili, Kouzarides T., Martignetti J.A.,
RA Peddie E., Satchwell S.C., Tomlinson P., Weston K.M., Bartrell B.G.;
RT "Analysis of the protein-coding content of the sequence of human
RT cytomegalovirus strain AD169.";
RT Curr. Top. Microbiol. Immunol. 154:125-169(1990).
CC -1- FUNCTION: PROBABLY INVOLVED IN DNA REPLICATION.
CC -1- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL8,
CC EHV-1 54, VZV 52 AND HCMV 102.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X17403; CAA35338.1; -
CC PIR: S09867; S09867.
DR PIR: S09867; S09867.
KW DNA replication.
SQ SEQUENCE 798 AA; 85613 MW; 6AA4E14B11E1F5EB CRC64;

Query Match 53.3%; Score 40; DB 1; Length 798;
Best local Similarity 60.0%; Pred. No. 39;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```

```

HEPA_HCMVA STANDARD; PRT; 798 AA.
ID HEPA_HCMVA
AC P16837;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)

```

```

DT 01-JUL-1993 (Rel. 26, Created)
ID HS2C_ARATH
AC HS2C_ARATH
ID P31170;
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Small heat shock protein, chloroplast precursor.
GN HSP21 OR AT4G27670 OR T29A15.160.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia; TISSUE=Leaf;
RC MEDLINE=91246118; PubMed=2038305;
RA Chen Q., Vierling E.;
RT "Analysis of conserved domains identifies a unique structural feature
RT of a chloroplast heat shock protein.";
RL Mol. Gen. Genet. 226:425-431(1991).
RN [2]
RP ERRATUM.
RC MEDLINE=91360086; PubMed=1886617;
RA Chen Q., Vierling E.;
RL Mol. Gen. Genet. 228:328-328(1991).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RC MEDLINE=93302705; PubMed=8391109;

```

RA OsterYoung K.W., Sundberg H., Vierling E.;  
 RT "Poly(A) tail length of a heat shock protein RNA is increased by  
 RT severe heat stress, but intron splicing is unaffected.";  
 RL Mol. Gen. Genet. 239:323-333(1993).  
 RM (4)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV, Columbia;  
 RX MEDLINE=20083488; PubMed=10617198;  
 RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,  
 RA Poll T., Duetschhoeft A., Stiekema W., Entlin K.D., Terryn N.,  
 RA Harris B., Ansoorge W., Brandt P., Grivell L.A., Rieger M.,  
 RA Weichselgartner M., de Simone V., Obermaier B., Maché R., Mueller M.,  
 RA Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidtheini T.,  
 RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,  
 RA Vos P., Hohnselt J., Zimmermann W., Wedler H., Ridley P.,  
 RA Langham S.-A., McCullagh B., Bilham L., Robben J.,  
 RA Van der Schueren J., Grymonprez B., Chung Y.-J., Vandenbussche F.,  
 RA Braeken M., Weltjens I., Voet M., Bastlaens I., Aert R., Defoor E.,  
 RA Weltzenegger T., Bothe G., Ramsperger U., Hilbert H., Braun M.,  
 RA Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,  
 RA Moolman P., Klein lankhorst R., Rose M., Hauf J., Koetter P.,  
 RA Beneliser S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,  
 RA De Keyser A., Blyssnaert C., Gielen J., Villarroel R., De Clercq R.,  
 RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,  
 RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McIay K., Mayes R.,  
 RA Petrelet A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,  
 RA Borkova D., Bloeker H., Scharfe M., Grimm M., Loehner T.-H.,  
 RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Aer S.,  
 RA Gabel C., Fuchs M., Faltmann B., Grandert K., Dauner D., Herzl A.,  
 RA Neumann S., Argitllo A., Vitale D., Liguori R., Piravandi E.,  
 RA Massenot O., Quigley F., Clabaud G., Muendlein A., Felber R.,  
 RA Schnabel S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,  
 RA Chiffois F., Cooke R., Berger C., Monfort A., Casacuberta E.,  
 RA Gibbons T., Weber N., Vandenbol M., Baquies M., Terol J., Torres A.,  
 RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacón D., Jesse F.,  
 RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Biele C.,  
 RA Frishman D., Haese D., Lemcke K., Mewes H.-W., Stocker S.,  
 RA Zaccarria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,  
 RA Parnell L., Dedha N., Gnoj L., Schütz K., Huang E., Spiegel L.,  
 RA Sekhon M., Murray J., Sheel P., Cordes M., Abu-Threideh J.,  
 RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,  
 RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,  
 RA Mink P., Bentley D., Fulton L., Miller N., Greco T., Kemp K.,  
 RA Kramer J., Fulton L., Martins E., Dante M., Pepin K., Hillier L.,  
 RA Nelson J., Speth J., Ryan E., Andrews S., Geisel C., Layman D.,  
 RA Du H., All J., Berghoff A., Jones K., Drone K., Cotton B., Joshi C.,  
 RA Antonolu B., Zidant M., Strong C., Sun H., Lamar B., Yordan C.,  
 RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,  
 RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Tili S.,  
 RA Grenat S., Shohdy N., Hasegawa A., Hamed A., Lohli M., Johnson A.,  
 RA Chen E., Marra M., Martienssen R., McCombie M.R.;  
 RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis  
 RT thaliana.";  
 RL Nature 402:769-777(1999).  
 CC -1- SUBCELLULAR LOCATION: Chloroplast.  
 CC -1- SIMILARITY: BELONGS TO THE SMALL HEAT SHOCK PROTEIN (HSP20)  
 CC FAMILY.  
 CC -1- SIMILARITY: PLANTS HAVE FOUR CLASSES OF HSP20. CLASS I AND II  
 CC WHICH ARE CYTOPLASMIC; CLASS III WHICH IS IN THE CHLOROPLAST  
 CC AND CLASS IV WHICH IS IN THE ENDOMEMBRANE. THIS PROTEIN BELONGS  
 CC TO CLASS III.  
 CC -1- This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X54102; CAA38036.1; -  
 CC DR EMBL: M94455; AAA32818.1; -  
 CC DR EMBL: AL161571; CAB81417.1; -

DR EMBL: AL035602; CAB838279.1; -  
 DR PIR: S16527; S16527.  
 DR PIR: S16005; S16005.  
 DR PIR: S35240; S35240.  
 DR InterPro: IPR002068; HSP20.  
 DR Pfam: PF00011; HSP20; 1.  
 DR PROSITE: PS01031; HSP20; 1.  
 KW Heat shock; Chloroplast; Transit peptide.  
 FT TRANSIT 1 19  
 FT CHAIN 20 227  
 FT SEQUENCE 227 AA; 25344 MW; B25400AF6C01972E CRC64;  
 SO  
 Query Match 52.7%; Score 39.5; DB 1; Length 227;  
 Best Local Similarity 62.5%; Pred. No. 12;  
 Matches 10; Conservative 3; Mismatches 2; Indels 1; Gaps 1;  
 QY 1 SSSIICSP-PSLSLS 15  
 DB 8 AASALCSPLAPSPSVS 23  
 RESULT 5  
 RUNT\_DROME STANDARD; PRT; 509 AA.  
 AC P22814;  
 DT 01-AUG-1991 (Rel. 19, Last Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Segmentation protein Runt.  
 GN Runt.  
 OS Drosophila melanogaster (Fruit fly).  
 CC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
 CC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
 CC Muscophora; Ephydroidea; Drosophilidae; Drosophila.  
 CC NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91065517; PubMed=2249771;  
 RA Kania M.A., Bonner A.S., Duffy J.B., Gergen J.P.;  
 RT "The Drosophila segmentation gene runt encodes a novel nuclear  
 RT regulatory protein that is also expressed in the developing nervous  
 RT system.";  
 RL Genes Dev. 4:1701-1713(1990).  
 RN [2]  
 RP SIMILARITY TO AML.  
 RX MEDLINE=92220161; PubMed=1560822;  
 RA Daga A., Tighe J.E., Calabi F.;  
 RT "Leukaemia/Drosophila homology.";  
 RL Nature 356:484-484(1992).  
 CC -1- FUNCTION: PLAYS A PIVOTAL ROLE IN REGULATING THE EXPRESSION OF  
 CC OTHER PAIR-RULE GENES SUCH AS EVE, FTZ, AND H.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- TISSUE SPECIFICITY: DEVELOPING CENTRAL AND PERIPHERAL NERVOUS  
 CC SYSTEM.  
 CC -1- DEVELOPMENTAL STAGE: MOST ABUNDANTLY EXPRESSED AT THE BLASTODERM  
 CC STAGE OF EMBRYOGENESIS.  
 CC -1- SIMILARITY: CONTAINS 1 RUNT DOMAIN.  
 CC -1- This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X56432; CAA39817.1; -  
 CC DR PIR: A36392; A36392.  
 CC HSSP: O60472; ICMD.  
 CC TRANSFAC: T01066;  
 CC FLYbase: FBgn0003300; run.  
 CC InterPro: IPR000040; AML1\_Runt.  
 CC Pfam: PF00853; Runt; 1.



CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: AF155157; AAD43327.1; -  
DR MGI: 1337130; Ap4b1.  
DR InterPro: IPR002553; Adaplin\_N.  
DR Pfam: PF01602; Adaplin\_N; 1.  
DR Coated pits; Endocytosis.  
SQ SEQUENCE 739 AA; 82618 MW; 1007D972BF8D0897 CRC64;  
  
Query Match 52.0%; Score 39; DB 1; Length 739;  
Best Local Similarity 58.8%; Pred. No. 53;  
Matches 10; Conservative 1; Mismatches 4; Indels 2; Gaps 1;  
  
OY 1 SSSILCSPL--PSLSLS 15  
Db : ||||| ||| ||  
529 TKQILCSKSPSPSLGLS 545  
  
RESULT 8  
YNO4\_YEAST STANDARD; PRT; 740 AA.  
ID YNO4\_YEAST  
AC P53907;  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DE Hypothetical 84.2 kDa protein in MFA2 intergenic region.  
GN YNL144C OR N1205 OR N1790.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=5288C;  
RX MEDLINE=96109932; PubMed=8619318;  
RA Maillet L., Bussereau F., Jacquet M.;  
RT "A 43.5 kb segment of yeast chromosome XIV, which contains MFA2,  
MFA2, CAP/SRY2, NAM9, FKBI/FPRI/RBP1, MOM22 and CPT1, predicts an  
RT adenosine deaminase gene and 14 new open reading frames.";  
RL Yeast 11:1195-1209(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=5288C / FY1679;  
RX MEDLINE=96287653; PubMed=8686380;  
RA Naar F., Becam A.-M., Herbert C.J.;  
RT "The sequence of 36.8 kb from the left arm of chromosome XIV reveals  
RT 24 complete open reading frames: 18 correspond to new genes, one of  
RT which encodes a protein similar to the human myotonic dystrophy  
RT kinase.";  
RL Yeast 12:169-175(1996).  
CC -1- SIMILARITY: TO YEAST YHR131C.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: Z46843; CAA86882.1; -  
DR EMBL: X92517; CAA63294.1; -  
DR EMBL: Z71420; CAA96027.1; -  
DR SGD: S0005088; YNL144C.  
KW Hypothetical protein.  
SQ SEQUENCE 740 AA; 84202 MW; 7A3D75999DC83D06 CRC64;

Query Match 52.0%; Score 39; DB 1; Length 740;  
Best Local Similarity 53.3%; Pred. No. 53;  
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
  
OY 1 SSSILCSPLPSLSLS 15  
Db : ||||| |||| : :  
42 SNSILDSPLPKVPE 56  
  
RESULT 9  
YNE6\_CAEEL STANDARD; PRT; 918 AA.  
ID YNE6\_CAEEL  
AC P30645;  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Probable 3',5'-cyclic phosphodiesterase R08D7.6 (EC 3.1.4.17).  
GN R08D7.6.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-Bristol N2;  
RX MEDLINE=94150718; PubMed=7906398;  
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,  
RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,  
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,  
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,  
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,  
RA Lareelle P., Lightning J., Lloyd C., Northmore B., O'Callaghan M.,  
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,  
RA Sims M., Smailton N., Smith A., Smith M., Sonhammer E., Staden R.,  
RA Sulston J., Thierry-Mieg J., Thomas K., Vaubin M., Vaughan K.,  
RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,  
RA Wohldman P.;  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
RT elegans.";  
RL Nature 368:32-38(1994).  
CC -1- CATALYTIC ACTIVITY: Nucleoside 3',5'-cyclic phosphate + H(2)O =  
CC nucleoside 5'-phosphate.  
CC -1- SIMILARITY: BELONGS TO THE CYCLIC NUCLEOTIDE PHOSPHODIESTERASE  
CC FAMILY. STRONG, TO MAMMALIAN TYPE 6 CGMP PHOSPHODIESTERASES.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: Z12017; CAA78052.1; -  
DR PIR: S24462; S24462.  
DR PIR: S41041; S41041.  
DR WormPep: R08D7.6; CE01047.  
DR InterPro: IPR003018; GAF.  
DR InterPro: IPR003607; ME\_PpIase\_HDC.  
DR InterPro: IPR002073; PDEase.  
DR Pfam: PF00233; PDEase; 1.  
DR Pfam: PF01590; GAF; 1.  
DR PRINTS: PRO0387; PDIESTERASE1.  
DR SMART: SM00065; GAF; 1.  
DR SMART: SM00471; HDC; 1.  
DR PROSITE: PS00126; PDEASE\_I; 1.  
KW Hypothetical protein; Hydrolyase; CGMP.  
SQ SEQUENCE 918 AA; 104238 MW; CA59910AF4CD6155 CRC64;  
  
Query Match 52.0%; Score 39; DB 1; Length 918;  
Best Local Similarity 72.7%; Pred. No. 67;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;







OX NCBI\_TaxId-9940;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC TISSUE-Pituitary;  
 RX MEDLINE-94059099; PubMed-7694577;  
 RA ILLING N., Jacobs G.F.M., Becker I.I., Flanagan C.A., Davidson J.S.,  
 RA Eales A., Zhou W., Sealton S.C., Millar R.P.;  
 RT "Comparative sequence analysis and functional characterization of the  
 RT cloned sheep gonadotropin-releasing hormone receptor reveal  
 RT differences in primary structure and ligand specificity among  
 RT mammalian receptors";  
 RL Biochem. Biophys. Res. Commun. 196:745-751(1993).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN-Scottish blackface; TISSUE-Pituitary;  
 RX MEDLINE-94040345; PubMed-8224516;  
 RA Brooks J., Taylor P.L., Saunders P.T.K., Eide K.A., Scruthers W.J.,  
 RA McNeilly A.S.;  
 RT "Cloning and sequencing of the sheep pituitary gonadotropin-releasing  
 RT hormone receptor and changes in expression of its mRNA during the  
 RT estrous cycle.";  
 RL Mol. Cell. Endocrinol. 94:R23-R27(1993).  
 RN [3]  
 RN SEQUENCE FROM N.A.  
 RC TISSUE-Liver;  
 RX MEDLINE-96235149; PubMed-8666259;  
 RA Campion C.E., Turzillo A.M., Clay C.M.;  
 RT "The gene encoding the ovine gonadotropin-releasing hormone (GNRH)  
 RT receptor: cloning and initial characterization.";  
 RL Gene 170:277-280(1996).  
 CC -1- FUNCTION: THIS RECEPTOR MEDIATES ITS ACTION BY ASSOCIATION WITH  
 CC G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND  
 CC MESSENGER SYSTEM.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>  
 CC or send an email to [license@sdb-sib.ch](mailto:license@sdb-sib.ch)).  
 CC -----  
 DR EMBL: L22215; AAC37336.1; -;  
 DR EMBL: X72088; CA50978.1; -;  
 DR EMBL: L42937; AAB38515.1; -;  
 DR EMBL: L43842; ABA41939.1; -;  
 DR EMBL: L43841; ABA41939.1; JOINED.  
 DR PIR: JN0882; JN0882;  
 DR InterPro: IPR000276; GPCR\_Rhodpsn.  
 DR Pfam: PF00001; 7tm\_1; 1.  
 DR PROSITE: PS00237; G\_PROTEIN\_RECEP\_F1\_1;  
 DR PROSITE: PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
 KW G-protein coupled receptor; Transmembrane; Glycoprotein.  
 FT DOMAIN 1  
 FT TRANSMEM 39 58  
 FT TRANSMEM 59 77  
 FT TRANSMEM 78 97  
 FT TRANSMEM 98 115  
 FT TRANSMEM 116 137  
 FT TRANSMEM 138 164  
 FT TRANSMEM 165 184  
 FT TRANSMEM 185 212  
 FT TRANSMEM 213 232  
 FT TRANSMEM 233 281  
 FT TRANSMEM 282 300  
 FT TRANSMEM 301 306  
 FT TRANSMEM 307 326  
 FT TRANSMEM 327 328  
 FT TRANSMEM 329 328  
 FT CARBOHYD 18 18  
 FT CARBOHYD 102 102  
 FT DISULFID 114 196  
 FT BY SIMILARITY.

FT CONFLICT 8 8 D -> N (IN REF. 2).  
 FT CONFLICT 27 27 S -> R (IN REF. 2).  
 FT CONFLICT 64 64 T -> A (IN REF. 2).  
 FT CONFLICT 312 312 A -> G (IN REF. 2).  
 SQ SEQUENCE 328 AA: 37684 MW: 0618374F33ECC6FE CRC64:  
 Query Match 50.7%; Score 38; DB 1; Length 328;  
 Best Local Similarity 58.8%; Pred. No. 32;  
 Matches 10; Conservative 4; Mismatches 1; Indels 2; Gaps 1;  
 QY 1 SSSIICSP--LPSSLS 15  
 DB 18 NSSILLPESLPETLRLS 34  
 RESULT 13  
 ID IRLS\_BURPS STANDARD; PRT; 464 AA.  
 AC 031396;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Sensor protein IRLS (EC 2.7.3.-).  
 GN IRLS.  
 OS Burkholderia pseudomallei (Pseudomonas pseudomallei).  
 OC Bacteria; Proteobacteria; beta subdivision; Burkholderia group;  
 OC Burkholderia.  
 OX NCBI\_TaxId-28450;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN-1026B;  
 RX MEDLINE-98053933; PubMed-9393784;  
 RA Jones A.L., Deshazer D., Woods D.E.;  
 RT "Identification and characterization of a two-component regulatory  
 RT system involved in invasion of eukaryotic cells and heavy-metal  
 RT resistance in Burkholderia pseudomallei";  
 RL Infect. Immun. 65:4972-4977(1997).  
 CC -1- FUNCTION: MEMBER OF THE TWO-COMPONENT REGULATORY SYSTEM IRLR/IRLS.  
 CC MAY BE INVOLVED IN INVASION OF EUKARYOTIC CELLS AND HEAVY-METAL  
 CC RESISTANCE. PROBABLY ACTIVATES IRLR BY PHOSPHORYLATION.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
 CC (Potential).  
 CC -1- SIMILARITY: CONTAINS 1 HISTIDINE KINASE DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>  
 CC or send an email to [license@sdb-sib.ch](mailto:license@sdb-sib.ch)).  
 CC -----  
 DR EMBL: AF005358; AAB92483.1; -;  
 DR InterPro: IPR003594; ATPBind\_ATPase.  
 DR InterPro: IPR004358; Bact\_sens\_pr\_C.  
 DR InterPro: IPR003660; HAMF.  
 DR InterPro: IPR004359; His\_Kin\_s1g.  
 DR InterPro: IPR003661; His\_Kin.  
 DR Pfam: PF00512; signal; 1.  
 DR Pfam: PF00672; HAMF; 1.  
 DR Pfam: PF02518; HATPase\_C; 1.  
 DR PRINTS: PR00344; BCTRLSENSOR.  
 DR SMART: SM00304; HAMF; 1.  
 DR SMART: SM00387; HATPase\_C; 1.  
 DR SMART: SM00388; HSKA; 1.  
 DR PROSITE: PS50109; HIS\_KIN; 1.  
 KW Sensory transduction; Transferrase; Kinase; Phosphorylation;  
 KW Transmembrane; Inner membrane; Zinc; Cadmium.  
 FT DOMAIN 1  
 FT TRANSMEM 14 33  
 FT TRANSMEM 35 166  
 FT TRANSMEM 167 187  
 FT TRANSMEM 188 464  
 FT PERIPLASMIC (POTENTIAL).

FT DOMAIN 249 463 HISTIDINE KINASE.  
 FT MOD\_RES 252 252 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 SQ SEQUENCE 464 AA; 50652 MW; 34FE275489AEDCBDCRC64;  
 Query Match 50.7%; Score 38; DB 1; Length 464;  
 Best Local Similarity 60.0%; Pred. No. 47;  
 Matches 9; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 Db 1 SSSIICSPPLPS 15  
 205 SSSRLAQLPELDT 219  
 RESULT 14  
 YCT3\_SCHPO STANDARD; PRT; 554 AA.  
 AC 059813; P78767; 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Putative amino-acid permease C794.03.  
 GN SPC794.03.  
 OS Schizosaccharomyces pombe (fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomyces.  
 NCBI\_TaxID=4896;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RX MEDLINE=21848401; PubMed=11859360;  
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgourou J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor A., Cronin A., Davis P., Felwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson A.,  
 RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Mount S., Mungall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rablinswitsch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volckaert G., Aert R., Robben J., Gymnopoulos B.,  
 RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Lebrach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Medler H., Wandut R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Diano S., Gloux S., Lelaure V., Mottier S.,  
 RA Gilbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Shpakovski G.V., Ussery D., Batteil B.G., Nurse P.;  
 RT The genome sequence of Schizosaccharomyces pombe.  
 RL Nature 415:871-880(2002).  
 [2]  
 RP SEQUENCE OF 198-554 FROM N.A.  
 RC STRAIN=PR745;  
 RX MEDLINE=98162722; PubMed=9501991;  
 RA Yoshida S., Kato K., Nakai K., Okayama H., Nojima H.;  
 RT Identification of open reading frames in Schizosaccharomyces pombe  
 RT cDNAs.  
 RL DNA Res. 4:363-369(1997).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1- SIMILARITY: BELONGS TO THE AMINO ACID PERMEASE FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: A023595; CA19131.1; -  
 DR EMBL: D89115; BA13777.1; -  
 DR InterPro: IPR002293; AA/rel\_pmease1.  
 DR InterPro: IPR004840; AAC-permease.  
 DR InterPro: IPR004841; Permease.  
 DR Pfam: PF00324; aa-permeases; 1.  
 DR PROSITE: PS00218; AMINO-ACID-PERMEASE\_1; FALSE\_NEG.  
 KW Hypothetical protein; Transport; Amino-acid transport; Transmembrane.  
 FT TRANSMEM 68 88  
 FT TRANSMEM 126 146 POTENTIAL.  
 FT TRANSMEM 176 196 POTENTIAL.  
 FT TRANSMEM 206 226 POTENTIAL.  
 FT TRANSMEM 288 308 POTENTIAL.  
 FT TRANSMEM 397 417 POTENTIAL.  
 FT TRANSMEM 492 512 POTENTIAL.  
 FT TRANSMEM 293 293 A -> D (IN REF. 2).  
 FT CONFLICT 293 308 VMEF -> RHVF (IN REF. 2).  
 FT CONFLICT 305 308 MISSING (IN REF. 2).  
 FT CONFLICT 313 313 A -> G (IN REF. 2).  
 FT CONFLICT 372 372 P -> A (IN REF. 2).  
 FT CONFLICT 412 412 A -> D (IN REF. 2).  
 FT CONFLICT 423 423  
 SQ SEQUENCE 554 AA; 60857 MW; F88E245437F1265C CRC64;  
 Query Match 50.7%; Score 38; DB 1; Length 554;  
 Best Local Similarity 63.6%; Pred. No. 57;  
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 Db 1 SSSIICSPPLPS 11  
 403 SALLCSNPLPS 413  
 RESULT 15  
 TRPX\_ARATH STANDARD; PRT; 621 AA.  
 ID TRPX\_ARATH.  
 AC P32069;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Anthranilate synthase component 1-2, chloroplast precursor  
 DE (EC 4.1.3.27).  
 GN ASA2 OR AT2G26960 OR T27A16.21.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Arabidopsida; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eucotsids II; Brassicales; Brassicaceae; Arabidopsis.  
 NCBI\_TaxID=3702;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93005721; PubMed=1392592;  
 RA Niyogi K.K., Fink G.R.;  
 RT "Two anthranilate synthase genes in Arabidopsis: defense-related  
 RT regulation of the tryptophan pathway";  
 RL Plant Cell 4:721-733(1992).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=20083487; PubMed=10617197;  
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,  
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,  
 RA Buell C.R., Ketchum K.A., Lee J.J., Renning C.M., Koo H.L.,  
 RA Moffat K.S., Cronin L.A., Shen M., Pail G., Van Aken S., Umayam L.,  
 RA Tallon L.J., Gill J.E., Adams M.D., Carreira A.J., Cressy T.H.,  
 RA Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,  
 RA Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,  
 RA Venter J.C.;  
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis  
 RT thaliana";  
 RL Nature 402:761-768(1999).  
 CC -1- CATALYTIC ACTIVITY: Chorismate + L-glutamine -> anthranilate +



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd

OM protein - protein search, using sw model

Run on: June 25, 2003, 17:13:30 ; Search time 1.93289 Seconds  
(without alignments)  
1705.612 Million cell updates/sec

Title: US-10-075-846-4\_COPY\_192\_207  
Perfect score: 75

Sequence: 1 SSSIICSPILPSLSLV 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

Database :

```

1:  sp.archaea:*
2:  sp.bacteria:*
3:  sp.fungi:*
4:  sp_human:*
5:  sp_invertebrate:*
6:  sp_mammal:*
7:  sp.mhc:*
8:  sp.organelle:*
9:  sp.phage:*
10: sp.plant:*
11: sp.podent:*
12: sp.virus:*
13: sp.yeast:*
14: sp_unclassified:*
15: sp_virus:*
16: sp.bacteriap:*
17: sp.archaeap:*

```

**Pred. No.** is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	49	65.3	490	10	Q8UY8	Q8UY8 arabidopsis
2	47	62.7	1298	11	Q8R46	Q8R46 mus musculus
3	46	61.3	94	16	Q9JMS9	Q9JMS9 escherichia
4	46	61.3	117	16	Q8X309	Q8X309 escherichia
5	46	61.3	490	10	Q9MAA0	Q9MAA0 arabidopsis
6	46	61.3	555	10	Q8W0Z7	Q8W0Z7 arabidopsis
7	46	61.3	724	11	Q912A3	Q912A3 mus musculus
8	46	61.3	724	11	Q922N3	Q922N3 mus musculus
9	44	58.7	887	4	Q1S134	Q1S134 homo sapien
10	43	57.3	82	9	Q8W6E1	Q8W6E1 bacterioph
11	43	57.3	138	5	Q9N7D6	Q9N7D6 leishmania
12	43	57.3	189	4	Q96EH4	Q96EH4 homo sapien
13	43	57.3	3567	11	Q9SE77	Q9SE77 mus musculus
14	42	56.0	113	11	Q9QWV3	Q9QWV3 mus musculus
15	42	56.0	114	17	Q9YBN2	Q9YBN2 atropyrum f
16	42	56.0	123	4	Q8WTW1	Q8WTW1 homo sapien

17	42	56.0	141	16	09Pht5	09Pht5 xylella fas
18	42	56.0	478	16	083294	083294 tireponema p
19	42	56.0	739	11	035613	035613 mus musculus
20	42	56.0	740	11	09QMT8	09QMT8 mus musculus
21	42	56.0	1010	1	09USN1	09USN1 schizosacch
22	41	54.7	153	2	09ADL7	09ADL7 polyangium
23	41	54.7	185	10	0857G3	0857G3 oryza sativ
24	41	54.7	216	6	09N223	09N223 oryza sativ
25	41	54.7	396	10	080871	080871 arabidopsis
26	41	54.7	542	4	09BRC0	09BRC0 homo sapien
27	41	54.7	608	3	08TPE4	08TPE4 gaeumannomy
28	41	54.7	922	10	09M497	09M497 arabidopsis
29	41	54.7	944	3	P87114	P87114 schizosacch
30	41	54.7	1449	10	09S749	09S749 arabidopsis
31	41	54.7	1449	4	094837	094837 homo sapien
32	41	54.7	2123	10	09LPEK3	09LPEK3 arabidopsis
33	40	53.3	91	12	08Q258	08Q258 kaikalur viru
34	40	53.3	91	12	08Q243	08Q243 sango virus
35	40	53.3	91	12	08QPEV6	08QPEV6 peaton viru
36	40	53.3	91	12	08QPEV4	08QPEV4 peaton viru
37	40	53.3	91	12	08QHN1	08QHN1 peaton viru
38	40	53.3	131	4	09P0F9	09P0F9 homo sapien
39	40	53.3	131	4	09P6P9	09P6P9 homo sapien
40	40	53.3	133	4	08WY82	08WY82 homo sapien
41	40	53.3	141	11	09ROE5	09ROE5 marmola mar
42	40	53.3	150	2	09F804	09F804 erwinia amy
43	40	53.3	247	5	062268	062268 caenorhabdi
44	40	53.3	254	3	09P960	09P960 aspergillus
45	40	53.3	262	16	084111	084111 chlamydia t

## ALIGNMENTS

RESULT 1									
09SUVR	ID	09SUVR	PRELIMINARY:	PRT:	490 AA.				
AC	09SUVR8:								
DT	01-MAY-2000	(TREMBLrel. 13, Created)							
DT	01-MAY-2000	(TREMBLrel. 13, Last sequence update)							
DT	01-DEC-2001	(TREMBLrel. 19, Last annotation update)							
DE	Hypothetical	53.0 kDa protein.							
GN	F4FL5_280.								
OS	Arabidopsis thaliana (Mouse-ear cress).								
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;								
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;								
OC	eucots II; Brassicales; Brassicaceae; Arabidopsis.								
OX	NCBI_TaxID=3702;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RA	Alcaraz J.P., Clabault G., Cottet A., Mache R., Mewes H.W., Lemcke K.,								
RA	Mayer K.F.X., Quetier F., Salenouat M.;								
RL	Submitted (APR-1999) to the EMBL/Genbank/DBJ databases.								
RP	[2]								
RP	SEQUENCE FROM N.A.								
RA	EU Arabidopsis sequencing project;								
RL	Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.								
DR	EMBL; AL049711; CAB41337.1; -.								
KW	Hypothetical protein.								
SO	SEQUENCE	490 AA;	53038 MW;	3784BD06F9580447C	CRC64;				
Query Match:									
	Best Local Similarity	65.3%;	Score 49;	DB 10;	Length 490;				
	Matches 11;	Conservative	1;	Mismatches 3;	Indels 0;	Gaps 0			
OY	1	SSSLICSPLPSTLS	15						
			:						
DB	129	SSSILMDVPLPLS	143						
RESULT 2									
08RAH6	ID	08RAH6	PRELIMINARY:	PRT:	1298 AA.				

AC 08R4H6;  
DT 01-JUN-2002 (TREMBlrel. 21, Created)  
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE Common site lymphoma/leukemia GEF.  
GN CLG.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BXN-2;  
RX PubMed-11839748;  
RA Himmel K.L., Bi F., Shen H., Jenkins N.A., Copeland N.G., Zheng Y.,  
RA Largeaspada D.A.;  
RT "Activation of Clg, a Novel DB1 Family Guanine Nucleotide Exchange  
RT Factor Gene, by Proviral Insertion at Evi24, a Common Integration Site  
RT in B Cell and Myeloid Leukemias.";  
RL J. Biol. Chem. 277:13463-13472(2002).  
DR EMBL: AF465238; AAL93134.1;  
SQ SEQUENCE 1298 AA; 139193 MW; 3451F4FC4539639C CRC64;  
Query Match 62.7%; Score 47; DB 11; Length 1298;  
Best Local Similarity 71.4%; Pred. No. 13;  
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
OY 3 SITCSPPLSLSLV 16  
DB 1048 SITCNPPLSLSDV 1061  
RESULT 3  
O9JMS9 PRELIMINARY; PRT; 94 AA.  
AC 09JMS9;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
DE YblB protein.  
GN YblB.  
OS Escherichia coli.  
OC Plasmid F.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Shimizu H., Saitoh Y., Suda Y., Uehara K., Sempel G., Mizobuchi K.;  
RT "Complete nucleotide sequence of the F plasmid: its implications for  
RT organization and diversification of plasmid genomes.";  
RT submitted (Apr-2000) to the EMBL/Genbank/DBJ databases.  
RL [2]  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-90317835; PubMed-2164585;  
RA Yoshioke Y., Fujita Y., Ohtsubo E.;  
RT "Nucleotide sequence of the promoter-distal region of the tra operon  
RT of plasmid R100, including traI (DNA helicase I) and tra genes.";  
RL J. Mol. Biol. 214:39-53(1990).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-87194554; PubMed-3032897;  
RA Saadi S., Maas W.K., Hill D.F., Bergquist P.L.;  
RT "Nucleotide sequence analysis of RepFIC, a basic replicon present in  
RT IncFII plasmids p307 and F, and its relation to the RepA replicon of  
RT IncFII plasmids.";  
RL J. Bacteriol. 169:1836-1846(1987).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-95337425; PubMed-7612932;  
RA Broom J.E., Hill D.F., Hughes G., Jones W.A., McNaughton J.C.,  
RA Stockwell P.A., Petersen G.B.;  
RT "Sequence of a transposon identified as Tn1000 (gamma delta).";

RL DNA Seq. 5:185-189(1995).  
RN [5]  
RP SEQUENCE FROM N.A.  
RA Eichenlaub R.;  
RT "F plasmid DNA complete mini-F region (F coordinates 40,301P to  
RT 49,869F).";  
RL Submitted (Jul-2000) to the EMBL/Genbank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-86139869; PubMed-3949712;  
RA Helberg M., Eichenlaub R.;  
RT "Twelve 43-base-pair repeats map in a cis-acting region essential for  
RT partition of plasmid mini-F.";  
RL J. Bacteriol. 165:1043-1045(1986).  
RN [7]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-99296678; PubMed-10366527;  
RA Manwaring N.P., Skurray R.A., Firth N.;  
RT "Nucleotide sequence of the F plasmid leading region.";  
RL Plasmid 41:219-225(1999).  
RN [8]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-94359430; PubMed-7915817;  
RA Frost L.S., Ippen-Ihler K., Skurray R.A.;  
RT "An analysis of the sequence and gene products of the transfer region  
RT of the F sex factor.";  
RL Microbiol. Rev. 58:162-210(1994).  
DR EMBL: AP001918; BAA97892.1;  
KW Plasmid; Complete proteome.  
SQ SEQUENCE 94 AA; 10260 MW; DF46050908F8BC0A CRC64;  
Query Match 61.3%; Score 46; DB 16; Length 94;  
Best Local Similarity 53.3%; Pred. No. 1.6;  
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
OY 2 SSTICSPPLSLSLV 16  
DB 12 TSITCSPMTSLKTSI 26  
RESULT 4  
O8X3O9 PRELIMINARY; PRT; 117 AA.  
AC 08X3O9;  
DT 01-MAR-2002 (TREMBlrel. 20, Created)  
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE Putative IS encoded protein.  
GN Z1827.  
OS Escherichia coli O157:H7.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=83334;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-O157:H7 / EDL933 / ATCC 700927;  
RX MEDLINE-21074935; PubMed-11206551;  
RA Perna N.T., Plunkett G., III, Burland V., Mau B., Glasner J.D.,  
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
RA Grofbeck E.J., Davis N.W., Lim A., Dimantanta E.T., Potamocous K.,  
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";  
RL Nature 409:529-533(2001).  
DR EMBL: AE005326; AAG5927.1;  
DR InterPro: IPR001207; Transposase\_mutatr.  
DR Pfam: PF00872; Transpo\_mutator; 1.  
KW Complete proteome.  
SQ SEQUENCE 117 AA; 13323 MW; 10A5292943909892 CRC64;  
Query Match 61.3%; Score 46; DB 16; Length 117;  
Best Local Similarity 53.3%; Pred. No. 2;

Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 SSILCSPPLSLSV 16  
:||:||||:| 1:  
Db 12 TSILCSPMTSLKTSI 26

## RESULT 5

Q9MAAO PRELIMINARY; PRT; 490 AA.  
AC Q9MAAO:  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
DE Putative branched-chain amino acid aminotransferase.  
GN T12H1.16.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. COLUMBIA;  
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B., Wu D.,  
RA Rongling C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,  
RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;  
RT "Arabidopsis thaliana chromosome III BAC T12H1 genomic sequence."  
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AC009177; AF27025.1; -;  
DR HSSP: P00510; IASG.  
DR InterPro: IPR001544; Amino-tran\_4.  
DR Pfam: PF01063; amino-tran\_4; 1.  
DR ProDom: PD001961; Amino-tran\_4; 1.  
KM Amino-transferase; Transferase.  
SQ SEQUENCE 490 AA; 54829 MW; D29B6876564B36BD CRC64;

Query Match 61.3%; Score 46; DB 10; Length 490;  
Best Local Similarity 68.8%; Pred. No. 7.7;  
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 SSILCSPPLSLSV 16  
|||:| |||| | 1:  
Db 210 SSILSSPLPPSLPV 225

## RESULT 6

Q8M0Z7 PRELIMINARY; PRT; 555 AA.  
AC Q8M0Z7:  
DT 01-MAR-2002 (TREMBlrel. 20, Created)  
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE Putative branched-chain amino acid aminotransferase  
DE (AT905190/T12H1.16).  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC Cheuk R., Chen H., Kim C.J., Meyers M.C., Banh J., Bowser L.,  
RA Carninci P., Chang E., Dale J.M., Goldsmith A.D., Hayashizaki Y.,  
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,  
RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,  
RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,  
RA Tang C.C., Toriumi M., Wu H.C., Yamada K., Yamamura Y., Yu G., Yu S.,  
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;  
RT "Arabidopsis cDNA clones."  
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.

RA Kim C.J., Chen H., Cheuk R., Meyers M.C., Shin P., Banh J.,  
RA Bowser L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,  
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,  
RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,  
RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,  
RA Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,  
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,  
RA Ecker J.R.;  
RT "Arabidopsis ORF clones."  
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF462849; AL158936.1; -;  
DR EMBL: AY090279; AL159040.1; -;  
DR InterPro: IPR001544; Amino-tran\_4.  
DR Pfam: PF01063; amino-tran\_4; 1.  
DR ProDom: PD001961; Amino-tran\_4; 1.  
KW Amino-transferase; Transferase.  
SQ SEQUENCE 555 AA; 62212 MW; 670A17B8D74B5D86 CRC64;

Query Match 61.3%; Score 46; DB 10; Length 555;  
Best Local Similarity 68.8%; Pred. No. 8.7;  
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 SSILCSPPLSLSV 16  
|||:| |||| | 1:  
Db 241 SSILSSPLPPSLPV 256

## RESULT 7

Q912A3 PRELIMINARY; PRT; 724 AA.  
AC Q912A3:  
DT 01-DEC-2001 (TREMBlrel. 19, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE Propionyl CoA-carboxylase alpha-subunit.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6;  
RA Miyazaki T., Ohura T., Kobayashi M.;  
RT "Fatal propionic acidemia in mice lacking propionyl-CoA carboxylase  
RT and its rescue by postnatal, liver-specific supplementation via a  
RT transgene."  
RL J. Biol. Chem. 0:0-0(2001).  
DR EMBL: AY046947; AL102364.1; -;  
DR InterPro: IPR001882; Biotin\_attach.  
DR InterPro: IPR000089; Biotin\_11poyl.  
DR InterPro: IPR000901; CPSase.  
DR Pfam: PF02785; Biotin\_carb\_C; 1.  
DR Pfam: PF00364; Biotin\_11poyl; 1.  
DR Pfam: PF00289; CPSase\_L-chain; 1.  
DR Pfam: PF02786; CPSase\_L-D2; 1.  
DR PROSITE: PS00188; BIOTIN; UNKNOWN\_1.  
DR PROSITE: PS00866; CPSASE\_1; UNKNOWN\_1.  
DR PROSITE: PS00867; CPSASE\_2; UNKNOWN\_1.  
SQ SEQUENCE 724 AA; 79660 MW; DB5EC4CBP40739FB CRC64;

Query Match 61.3%; Score 46; DB 11; Length 724;  
Best Local Similarity 50.0%; Pred. No. 11;  
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 SSILCSPPLSLSV 16  
:|| ||||| : 1:  
Db 654 TSILCSPMPCVVAV 669

## RESULT 8

Q922N3 PRELIMINARY; PRT; 724 AA.  
ID Q922N3  
AC Q922N3;

```
DR 01-DEC-2001 (TREMBlrel. 19, Created)
DR 01-DEC-2001 (TREMBlrel. 19, last sequence update)
DR 01-JUN-2002 (TREMBlrel. 21, last annotation update)
DE Similar to propionyl coenzyme A carboxylase, alpha polypeptide.
OS Mus musculus (mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; BC006915; AAB06915.1;
DR InterPro: IPR001882; Biotin_atlacth.
DR InterPro: IPR000089; Biotin_lipoyl.
DR InterPro: IPR009001; CPSase.
DR Pfam: PF02785; Biotin_carb_C; 1.
DR Pfam: PF00364; Biotin_lipoyl; 1.
DR Pfam: PF00289; CPSase_L-chain; 1.
DR Pfam: PF02786; CPSase_L_D2; 1.
DR PROSITE: PS00188; BIOTIN; UNKNOWN_1.
DR PROSITE: PS00866; CPSASE_1; UNKNOWN_1.
DR PROSITE: PS00867; CPSASE_2; UNKNOWN_1.
SQ SEQUENCE 724 AA; 79940 MW; 8B646617F9E399A2 CRC64;

Query Match
Best Local Similarity 61.3%; Score 46; DB 11; Length 724;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 SSSILCSPLPSLSV 16
DB 654 TSSTLCSMPGVVAV 669

RESULT 9
ID Q15134 PRELIMINARY; PRT; 887 AA.
AC Q15134;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)
DE Phosphatidylinositol 3-kinase.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE OF 1-280 FROM N.A.
RX MEDLINE-95354652; PubMed-7628435;
RA Volinia S., Dhand R., Vanhesebroeck B., MacDougall L.K.,
RA Zvelebil M.J., Domin J., Panaretou C., Waterfield M.D.;
RT "A human phosphatidylinositol 3-kinase complex related to the yeast
RT Vps34p-Vps15p protein sorting system.";
RL EMO J. 14:3339-3348(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA Volinia S.;
RL Submitted (DEC-1994) to the EMBL/Genbank/DBJ databases.
DR EMBL; Z46973; CAA87094.1;
DR InterPro: IPR000008; C2.
DR InterPro: IPR001263; PI3K.
DR InterPro: IPR002420; PI3K_C2.
DR InterPro: IPR000403; PI3_P14_kinase.
DR Pfam: PF00613; PI3Ka; 1.
DR Pfam: PF00792; PI3K_C2; 1.
DR Pfam: PF00454; PI3_P14_kinase; 1.
DR SMART; SM00239; C2; 1.
DR SMART; SM00145; PI3Ka; 1.
DR SMART; SM00146; PI3Kc; 1.
DR SMART; SM00142; PI3K_C2; 1.
DR PROSITE: PS00915; PI3_4_KINASE_1; 1.
DR PROSITE: PS00916; PI3_4_KINASE_2; 1.
DR PROSITE: PS00290; PI3_4_KINASE_3; 1.
```

```
KW Kinase.
SQ SEQUENCE 887 AA; 100988 MW; 2078660E7D9FEF187 CRC64;

Query Match
Best Local Similarity 58.7%; Score 44; DB 4; Length 887;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SSSILCSPLPSLS 13
DB 442 SSQITSPSPSVS 454

RESULT 10
ID Q8M6E1 PRELIMINARY; PRT; 82 AA.
AC Q8M6E1;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, last annotation update)
DE Putative major coat protein.
OC Bacteriophage VSKK.
OC Viruses; ssDNA viruses; Inoviridae.
OX NCBI_TaxID=180503;
RN [1]
RP SEQUENCE FROM N.A.
RA Sinha K.K., Ghosh A.;
RT "Complete nucleotide sequence of single-stranded filamentous
RT vibriophage VSKK.";
RL Submitted (NOV-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF452449; AAL40839.1;
KW Coat protein.
SQ SEQUENCE 82 AA; 8804 MW; 62BE0B7157F62D86 CRC64;

Query Match
Best Local Similarity 57.3%; Score 43; DB 9; Length 82;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 3 SSSILCSPLPSLS 13
DB 16 SVICPMPMSMS 26

RESULT 11
ID Q9N7D6 PRELIMINARY; PRT; 138 AA.
AC Q9N7D6;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, last annotation update)
DE Possible serine/threonine kinase pkn8 (Fragment).
GN LM28-292.
OC Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-FRIEDLIN;
RA Murphy L., Quail M., Harris D., Rajandream M., Ivens A., Barrell B.;
RL Submitted (JUL-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AL390935; CAC01001.1;
DR InterPro: IPR000719; Euk_pkinase.
DR ProDom: PD000001; Euk_pkinase; 1.
KW Kinase.
FT NON_TER
FT NON_TER
SQ SEQUENCE 138 AA; 15163 MW; 2AF74AF488B12D6 CRC64;

Query Match
Best Local Similarity 57.3%; Score 43; DB 5; Length 138;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 SSSILCSPLPSLS 15
DB 11111111111111111111
```



DB 120 SSSICSPPLPSLSL 134

# RESULT 12

096EH4 PRELIMINARY: PRT: 189 AA.

AC 096EH4; 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)  
 DE Similar to RIKEN CDNA 4933430F08 gene.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LUNG;  
 RA Strausberg R.;  
 RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC012330; AAH12330.1; -  
 SQ SEQUENCE 189 AA; 19503 MW; BC1D0F5B688D62E7 CRC64;

Query Match 57.3%; Score 43; DB 4; Length 189;  
 Best Local Similarity 71.4%; Pred. No. 10;  
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 SSSICSPPLPSLSL 14  
 DB 60 SPSLQCPPLPSLSL 73

# RESULT 13

09ES77 PRELIMINARY: PRT: 3567 AA.

AC 09ES77; 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)  
 DE Polydom protein precursor.  
 GN POLYDOM.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C3H/HEUSIC;  
 RX MEDLINE=20517255; PubMed=11062057;  
 RA Gilgès D., Vinit M.-A., Callebaut I., Coulombel L., Cacheux V.,  
 RA Romeo P.-H., Vigon I.;  
 RT "Polydom : a secreted protein with pentraxin, complement control  
 RT protein, epidermal growth factor and von Willebrand factor A  
 RT domains";  
 RL Biochem. J. 352:49-59 (2000).  
 DR EMBL; AF206329; AAG32160.1; -  
 DR HSSP; P00740; LEDM.  
 DR MGD; MGI:1928849; Polydom.  
 DR InterPro: IPR000152; Asx\_hydroxyl.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR000742; EGF-2.  
 DR InterPro: IPR001881; EGF\_CA.  
 DR InterPro: IPR001438; EGF-II.  
 DR InterPro: IPR003410; Hyalin.  
 DR InterPro: IPR001759; Pentaxin.  
 DR InterPro: IPR000436; Sushl\_SCR\_CCP.  
 DR InterPro: IPR002035; WVF\_A.  
 DR Pfam; PF00008; EGF; 10.  
 DR Pfam; PF02494; HXR; 2.  
 DR Pfam; PF00354; pentaxin; 1.  
 DR Pfam; PF00084; sushl; 33.  
 DR Pfam; PF00092; wvf; 1.  
 DR PRINTS; PR00010; EGFBL00D.

DR PRINTS; PR00895; PENTAXIN.  
 DR PRINTS; PR00453; VWFADOMAIN.  
 DR Prodom; PD002153; Pentaxin; 1.  
 DR SMART; SM00032; CCP; 34.  
 DR SMART; SM00181; EGF; 15.  
 DR SMART; SM00179; EGF\_CA; 9.  
 DR SMART; SM00001; EGF-like; 3.  
 DR SMART; SM00159; PTX; 1.  
 DR SMART; SM00327; WVF; 1.  
 DR PROSITE; PS00010; ASX\_HYDROXYL; 5.  
 DR PROSITE; PS00022; EGF\_1; UNKNOWN\_9.  
 DR PROSITE; PS01186; EGF-2; 11.  
 DR PROSITE; PS01187; EGF\_CA; 6.  
 DR PROSITE; PS0234; WVF; 1.  
 DR Signal.  
 KW SIGNAL.  
 FT SIGNAL.  
 SQ SEQUENCE 3567 AA; 387391 MW; 6FBA6276E12293B5 CRC64;

Query Match 57.3%; Score 43; DB 11; Length 3567;  
 Best Local Similarity 62.5%; Pred. No. 1.6e+02;  
 Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 1 SSSICSPPLPSLSY 16  
 DB 1622 SSMFCSDCPDLEGSV 1637

# RESULT 14

090WV3 PRELIMINARY: PRT: 113 AA.

AC 090WV3; 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)  
 DE Fas-binding protein Daxx (Fragment).  
 GN DAXX OR DAXX.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129SVJ;  
 RA Rowen L., Olin S., Madan A., Loretz C., James R., Dors M., Mix L.,  
 RA Hall J., Lasky S., Hood L.;  
 RT "Sequence of the mouse major histocompatibility locus class II  
 RT region";  
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF100956; AAC69891.1; -  
 DR MGD; MGI:1197015; Daxx.  
 DR InterPro: IPR005012; Daxx.  
 DR Pfam; PF03344; Daxx; 1.  
 DR NON\_TER  
 FT NON\_TER  
 SQ SEQUENCE 113 AA; 12086 MW; DA4E6A725336A511 CRC64;

Query Match 56.0%; Score 42; DB 11; Length 113;  
 Best Local Similarity 69.2%; Pred. No. 9.5;  
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 SSSICSPPLPSLSL 14  
 DB 68 TSSLCSPSPSL 80

# RESULT 15

09YBN2 PRELIMINARY: PRT: 114 AA.

AC 09YBN2; 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)  
 DE Hypothetical protein APE1566.  
 GN APE1566.

OS Aeropyrum pernix.  
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcates;  
OC Desulfurococcaceae; Aeropyrum.  
OX NCBI\_TaxID=56636;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-K1;  
RX MEDLINE-93310339; PubMed-10382966;  
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Halkawa Y.,  
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Anai A., Kosugi H.,  
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,  
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,  
RA Yamazaki J., Kishida N., Oguchi A., Aoki K.-I., Kubota K.,  
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;  
RT \*Complete genome sequence of an aerobic hyper-thermophilic  
RT crenarchaeon, Aeropyrum pernix K1.";  
RL DNA Res. 6:83-101(1999).  
DR EMBL; AP000062; BAA80566.1; -;  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 114 AA; 11565 MW; 130EEE20695FA670 CRC64;

Query Match 56.0%; Score 42; DB 17; Length 114;  
Best Local Similarity 54.5%; Pred. No. 9.6;  
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 6 CSPPLSLSLV 16  
I:|:|:|:  
Db 28 CTPPSMNLST 38

Search completed: June 25, 2003, 17:17:00  
Job time : 3.93289 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2003, 20:59:10 ; Search time 107 Seconds  
(without alignments)  
4700.466 Million cell updates/sec

Title: US-10-075-846-3  
Perfect score: 1640  
Sequence: 1 atgacacactgttcctgc.....taaaaaaaaaaaaaaaaaa 1640

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_NA: \*  
1: /cgn2\_6/ptodata/1/ina/5A-COMB.seq: \*  
2: /cgn2\_6/ptodata/1/ina/5B-COMB.seq: \*  
3: /cgn2\_6/ptodata/1/ina/6A-COMB.seq: \*  
4: /cgn2\_6/ptodata/1/ina/6B-COMB.seq: \*  
5: /cgn2\_6/ptodata/1/ina/PCTUS-COMB.seq: \*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed.  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	193	11.8	1555	US-08-809-802-11	Sequence 11, Appl
2	187	11.4	4621	US-09-592-891A-13	Sequence 13, Appl
3	182.2	11.1	2310	US-08-417-330A-15	Sequence 15, Appl
4	181.6	11.1	1609	US-09-592-891A-11	Sequence 11, Appl
5	179	10.9	3958	US-08-435-933-5	Sequence 5, Appl
6	179	10.9	3958	PCT-US96-06035-5	Sequence 5, Appl
7	174.2	10.6	1640	US-09-592-891A-12	Sequence 12, Appl
8	168	10.2	1368	US-09-130-339-1	Sequence 1, Appl
9	161.2	9.8	1297	US-09-130-339-3	Sequence 3, Appl
10	157.8	9.6	1657	US-08-554-659-3	Sequence 3, Appl
11	157.8	9.6	1844	US-09-002-361-1	Sequence 1, Appl
12	156.4	9.5	1866	US-08-417-330A-19	Sequence 19, Appl
13	156.2	9.5	1657	US-08-554-659-1	Sequence 1, Appl
14	156.2	9.5	1732	US-08-417-330A-17	Sequence 17, Appl
15	146.6	8.9	669	US-09-002-361-7	Sequence 7, Appl
16	141.8	8.6	1491	US-08-137-614A-3	Sequence 3, Appl
17	141.8	8.6	1491	US-08-768-301-1	Sequence 1, Appl
18	138.6	8.5	1638	US-08-417-330A-13	Sequence 13, Appl
19	137.6	8.4	1519	US-09-002-361-4	Sequence 4, Appl
20	135.6	8.3	2191	US-08-417-330A-11	Sequence 11, Appl
21	135.2	8.2	1052	US-09-592-891A-8	Sequence 8, Appl
22	131.6	8.0	2066	US-08-072-064-7	Sequence 7, Appl
23	130	7.9	2066	US-08-072-064-2	Sequence 2, Appl
24	128.4	7.8	2066	US-08-072-064-3	Sequence 3, Appl
25	128.4	7.8	2066	US-08-072-064-5	Sequence 5, Appl
26	125.8	7.7	1650	US-08-459-100A-1	Sequence 1, Appl
27	125.8	7.7	1650	PCT-US94-09589-1	Sequence 1, Appl

28	124	7.6	1707	US-08-809-802-7	Sequence 7, Appl
29	119.6	7.3	1970	US-08-072-064-9	Sequence 9, Appl
30	119.6	7.3	1970	PCT-US92-08558-3	Sequence 3, Appl
31	119.6	7.3	2059	PCT-US92-08558-2	Sequence 2, Appl
32	111.2	6.8	1542	US-08-249-112-1	Sequence 1, Appl
33	111.2	6.8	1542	PCT-US95-06556-1	Sequence 1, Appl
34	109	6.6	2635	US-08-768-301-3	Sequence 3, Appl
35	106.2	6.5	2508	US-09-627-650B-8	Sequence 8, Appl
36	106.2	6.5	2508	US-09-436-063C-8	Sequence 8, Appl
37	106.2	6.5	2544	US-09-627-650B-4	Sequence 4, Appl
38	106.2	6.5	2544	US-09-436-063C-4	Sequence 4, Appl
39	106.2	6.5	2601	US-09-627-650B-10	Sequence 10, Appl
40	106.2	6.5	2601	US-09-436-063C-10	Sequence 10, Appl
41	102.8	6.3	762	US-09-592-891A-7	Sequence 7, Appl
42	101.6	6.2	1917	US-09-627-650B-6	Sequence 6, Appl
43	101.6	6.2	1917	US-09-436-063C-6	Sequence 6, Appl
44	97.8	6.0	1479	US-08-249-112-2	Sequence 2, Appl
45	97.8	6.0	1479	PCT-US95-06556-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1  
US-08-809-802-11  
Sequence 11, Application US/08809802  
Patent No. 6455276  
GENERAL INFORMATION:  
APPLICANT: Le Bourdellies, Beatrice  
APPLICANT: Whiting, Paul John  
TITLE OF INVENTION: HUMAN ALPHA 4 RECEPTOR SUBUNIT  
TITLE OF INVENTION: OF THE GABA-A RECEPTOR  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Merck & Co., Inc.  
STREET: P.O. Box 2000, 126 E. Lincoln Ave.  
CITY: Rahway  
STATE: NJ  
COUNTRY: USA  
ZIP: 07065-0900  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows  
SOFTWARE: FASTED for Windows Version 2.0b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/809, 802  
FILING DATE: 19-JUN-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB95/02323  
FILING DATE: 29-SEP-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Yablonsky, Michael D  
REGISTRATION NUMBER: 40,407  
REFERENCE/DOCKET NUMBER: T1292  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 732-594-4678  
TELEFAX: 732-594-4720  
TELEX:  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1555 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: Coding Sequence  
LOCATION: 47...1402  
OTHER INFORMATION:  
US-08-809-802-11

Query Match 11.8%; Score 193; DB 4; Length 1555;  
 Best Local Similarity 53.3%; Pred. No. 6.3e-47;  
 Matches 486; Conservative 0; Mismatches 380; Indels 45; Gaps 2;

142 CTGACAACTTATGGGGGACATCTGATATGATGAGGAGATGAGCCCAATTTTAA 201  
 167 CTGCCCCACCTGGACGGGCTGATAGCCGGTTAGCCCGCACTTCGGGCTGGCATCGGA 226  
 202 GAGCCACCCGTGAAGCTGACCTGCAACATCTTCATCAACAGTTTCAGCTCCGACCAAG 261  
 227 GGGCCCCCGGTGAATGGGCCCTTGGCCCTGGAGTGGCCAGCATCAACCATCTCAGAG 286  
 262 ACCACATATGACATCCGGGATGCTCTTCTTGGGCAACAGTGAAGATCCACCGCTG 321  
 287 GCCAACATGGAGTACACCATGACGAGTGTCTCTGACACAGCTGGGGGACACAGGCTC 346  
 322 TCCATCCGATATATCTGATGATCTGATCTGACCTGATCCCTCATGCTGATCTATC 381  
 347 TCCATC---AACACACCAACGAGACCTTGGCTGGACAGCCGCTTCGTGACAAAGCTG 403  
 382 TGGAGCCGACCTCTTCTTCTTATGAGAAAGGGCCCACTTCATGAGGTGACCAAG 441  
 404 TGGGTGCGCGACACCTTCATCTGTAAGCAGCAAGTGGCTGTCCAGAGTGCAGTGC 463  
 442 GACACAAAGTTACTGCGCATCTTCAAGATGGGAATGCTGTACAGCATCAGGCTGAC 501  
 464 GAGAACAGCTCACTCCGCTGACAGCCGACGGGGTATCTGTACACATCCGATCAC 523  
 502 CTCATTTTGTCTCTGATGAGATGAGATCCCAAGTTCCTCATGACATCAGAGCTGACG 561  
 524 TCCACTGTGGCTGCGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 583  
 562 ATGACGCTTGAAGCTATCTACTCTGACAGCCCTGACATCTGTCTACTTCAAGTT 621  
 584 CTGACCTGAGAGCT-----AC 601  
 622 GGCATACCATGAGAAAGCTGCTTGTGAGAGGCTGGAGATGCTCTGCTGCTCAAGTG 681  
 602 GGTACTCATGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 661  
 682 GCTGAGGGGCTGACTCTGCCCCAGTTTATCTTGGGGATGAGAAAGATCTAGGCTGTT 741  
 662 CTGGAACAGCTGAGAGCTGGCGAGTTCACATCAACAGCTACCGCTTCACCAAGAGCTG 721  
 742 ACCAAGCACTACACACAGGAAATTCACCTGATCGAGTAAAGTTTCACTGGAACG 801  
 722 ATGAACCTCAAGTCCGCTGGCCAGTTCACAGGCTGACCTGCACTTCCACCTGGAGG 781  
 802 CAGATGGGCTACTATCTGATCAATGATCAATCCCAAGCTACTGATCTGATCTGCTG 861  
 782 AACCGCGGCTGTATCATATCAATCAATCAATGCTGCTGCTGCTGCTGCTGCTGCTG 841  
 862 TGGGTCTCTTCTGATCAACATGATGATGATGATGATGATGATGATGATGATGATGAT 921  
 842 TGGGTCTCTTCTGATCAACATGATGATGATGATGATGATGATGATGATGATGATGAT 901  
 922 ACCGTCTCTCTGATCAACATGATGATGATGATGATGATGATGATGATGATGATGAT 981  
 902 ACGGTCTCTCTGATCAACATGATGATGATGATGATGATGATGATGATGATGATGAT 961  
 982 TACGTGAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1041  
 962 GCCATCAAGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1021  
 1042 CTGAGATATGC 1052  
 1022 GTGAGATAGC 1032

RESULT 2  
 US-09-592-891A-13  
 ; Sequence 13, Application US/09592891A  
 ; Patent No. 6329174

GENERAL INFORMATION:  
 APPLICANT: Xiao-Zhou Michelle Wang  
 APPLICANT: Xavier Georges Sarda  
 APPLICANT: Michael David Tomalski  
 APPLICANT: Vincent Paul Mary Wingate  
 TITLE OF INVENTION: Helicobacter Glutamate Receptor  
 FILE REFERENCE: A32815 072667.0118  
 CURRENT APPLICATION NUMBER: US/09/592,891A  
 NUMBER OF SEQ ID NOS: 14  
 SOFTWARE: FastSeq for Windows Version 3.0  
 SEQ ID NO 13  
 TYPE: DNA  
 ORGANISM: Helicobacter virescens  
 FEATURE:  
 NAME/KEY: unsure  
 LOCATION: (764)...(764)  
 OTHER INFORMATION: n is a, t, y, or c  
 US-09-592-891A-13

Query Match 11.4%; Score 187; DB 4; Length 4621;  
 Best Local Similarity 54.3%; Pred. No. 6.9e-45;  
 Matches 472; Conservative 0; Mismatches 346; Indels 48; Gaps 3;

208 CCCGTGAAGCTGACCTGACATCTTCATCAAGTTTCAAGTCCGTCACCAAGACCA 267  
 330 CCAGCGTATAGGCTGACCTATATTTGTCGAAGTATCAAAAGATCATGACGTACCA 389  
 268 ATGAGCTACCGGGAATGCTTCTTGGCAACAGTGAATATCAACAGCCGTCTCTAC 327  
 390 ATGGAATACCTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 449  
 328 CGAGATATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 387  
 450 AATAATCTGAGAGTCCGCTCAATATCTGACACTGACCGAAGCAACAGTCTGATG 509  
 388 CCAGACCTCTTCTTGTATGAGAAAGGGCCCACTTCATGAGTGCACAGGACCAAC 447  
 510 CCTGATCATCTTCTTCAACAGAGAAAGGATCTTTCACCAACATCATCATCTCCGAC 569  
 448 AAGTACTGCGCATCTTCAAGATGAGATGATGATGATGATGATGATGATGATGATGAT 507  
 570 GTGATACCTCGGATCTTCCCAAGGCAACGCTGTGACATCCGAATCTCTTGACG 629  
 508 TTGCTGCTGATGAGACCTCAAGAACTTCCCATGAGCATCAAGAGCTGACAGATCAG 567  
 630 CTCGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 689  
 568 CTGAGACCTCATCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 627  
 690 ATGGCTACTT-----ATGGTTGG 707  
 628 ACCATGAAGACCTGCTTGTGAGTGGCTGGAAGATGCTGCTGCTGCTGCTGCTGCTGAG 687  
 708 ACCACGAGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 764  
 688 GGGGTGACTGCGCCCATTTATCTTGGGGATGAGAAAGATCAAGCTGTTGATCAAG 747  
 765 AACTTACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 821  
 748 CACTACACACAGGAAATTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 807  
 822 AAGACTAATACCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 881  
 808 GGTACTATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 867  
 882 AGTACTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 941  
 868 TCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 927  
 942 TCTTCTGCTGACAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1001



```

: CURRENT APPLICATION NUMBER: US/09/592,891A
: CURRENT FILING DATE: 2001-01-22
: NUMBER OF SEQ ID NOS: 14
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 11
: LENGTH: 1609
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: PCR amplified fragment of Helicobacter DNA cloned
: OTHER INFORMATION: Into pCR2.1-TOPO vector (Invitrogen)
US-09-592-891A-11
```

```

Query Match      11.3%; Score 181.6; DB 4; Length 1609;
Best Local Similarity 54.2%; Pred. No. 1.5e-43;
Matches 469; Conservative 0; Mismatches 349; Indels 48; Gaps 3;
```

```

OY 208 CCCGTGACGCTGACGCTGCAACATCTTCATCAACAGTTTCAGCTCCGTCACCAAGACCACA 267
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 334 CCAGCGGTAGTGGCGGTGCAATATATTTGTCGAAGTATATCAAAAGATGATGACGTGAC 393
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 268 ATGAGTACCGGGGATGATGCTCTGCGGCAACAGGATGACCAAGCCGCTGCTAC 327
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 394 ATGGAATACCTCCGTACATTAACGTTTCGGGAACAATGTTAGTGAACGCTCAAAATTC 453
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 328 CGAAGATATCTGATGACCTCTGACCTGATCCCTCATGCTGACCTATCTGAGAG 387
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 454 AATAATCTGGAGTGCCCTCAAAATACCTGACACTGACTGAAGCAACAGAGTGTGATG 513
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 388 CGAGACTCTTTTGTGTAATGAGAAAGGGCCCACTTCATGAGGTGACCAAGGACAC 447
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 514 CTTGATCTATCTCTCCCAAGAGAGAGGATTCATTCACAAATCATCATGCGGAC 573
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 448 AAGTACTGCGCATCTCAAGATGGGATGCTGTGACAGTACAGCTGACCTGACCTCAT 507
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 574 GTGTACATCCGATCTTCCCAACAGGACGTGCTGTACAGCATCCGATCTCCCTGACG 633
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 508 TTGCTCTGCTGATGACCTCAAGAACTCCCATGACATCCAGAGTGCAGAGTACAG 567
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 634 CTCTCGTGGCCCATGAACTCAAGTTGTACCCCTGATTAAGACAGACTCTGCTCAGG 693
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 568 CTTGAGAGCTCATCACTGACCTGAGCCCTGCTGCACTCTGCTGATTTAGTTGGTAC 627
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 694 ATGGCTAGTT-----ATGGTGG 711
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 628 ACCATGAAGACCTCGTGTGATGAGTGTGAGAGTCTCTGCTGTCAAGTGGCTGAG 687
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 712 ACCACAGACGATAGTTCATGATGAGAGAGGAGCAGC--CGTGCAGAGTGGTGA 768
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 688 GGGCTGACTGCGCCCAAGTTATCTGCGGAGTGAAGAGATCTAGGCTGTTTACCAG 747
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 769 AACTTACACTGCTCGCTGCTCAGC--GCTGAGAGAGTTCTCACTGCTGCTCAACAGT 825
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 748 CACTACACACAGAGGAAATTCACCTGATGAGGTAAGTTTCAACGAGGAGGACAGT 807
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 826 AAGACTAATACCGGTGATACAGTTGCTGATGAGGTACCTGCTTTCAACGAGGATTC 885
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 808 GGGTACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 867
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 886 AGTTACTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 945
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 868 TCTTCTGATCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 927
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 946 TCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1005
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 928 CTGACATGACACACAGAGAGTCTGCTCCGGGCTCTTTGCTTAAGGTGTCTGACGTG 987
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1006 CTGACATGACACACAGAGAGTCTGCTCCGGGCTCTTTGCTTAAGGTGTCTGACGTG 1065
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 988 AAGCAATGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1047
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1066 AAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1125
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

OY 1048 TATGCTGCAATATATTTGTTCTCG 1073
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1126 TTTGCGCTGCTCAACTATGCTGCTCG 1151
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

## RESULT 5

```

US-08-435-933-5
: Sequence 5, Application US/08435933
: Patent No. 5693492
: GENERAL INFORMATION:
: APPLICANT: Gully, Doris F.
: APPLICANT: Arena, Joseph P.
: APPLICANT: Paresa, Philip S.
: APPLICANT: Liu, Ken K.
: TITLE OF INVENTION: DNA ENCODING GLUTAMATE GATED CHLORIDE
: CHANNELS
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: John W. Wallen III
: STREET: 126 East Lincoln Avenue
: CITY: Rahway
: STATE: New Jersey
: COUNTRY: US
: ZIP: 07065-0907
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/435,933
: FILING DATE:
: CLASSIFICATION: 514
: ATTORNEY/AGENT INFORMATION:
: NAME: Wallen, III John W.
: REGISTRATION NUMBER: 35,403
: REFERENCE/DOCKET NUMBER: 19264
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (908) 594-3905
: TELEFAX: (908) 594-4720
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3958 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
US-08-435-933-5
```

```

Query Match      10.9%; Score 179; DB 1; Length 3958;
Best Local Similarity 53.7%; Pred. No. 1.4e-42;
Matches 473; Conservative 0; Mismatches 360; Indels 48; Gaps 3;
```

```

OY 193 AATTTTAAAGCCGCCACCGCTGACAGTGTGACATCTTCATCAACAGTTTCACTCC 252
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 570 AATGGCACAATGATGCTCCCGCATGATGATGAATTAATCTGTCGACATATATGAGC 629
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 253 GTACCAAGACCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 312
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 630 ATTAGTATATTTAAATGAGTACAGTGTCACTTAACCTTCCTGTAACAGTGTGACGAT 689
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 313 CCAGCGCTGCTCAACGGAATATCTGATGATGATGATGATGATGATGATGATGATGATG 372
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 690 GAACGCTCAAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 749
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 373 GACTTATCTGGAACCAACCTCTTTTGTGTAATGAGAGAGGAGGACATTCATGAG 432
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 750 AACCGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 809
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 433 GTACCAAGGACCAACAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 492
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 810 ATCATCATGCCCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 869
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

OY 493 AGGTCAGCCCTCATTTTGTCTCTGCTGATGAGCACTCAAGAACTTCCCATGAGACATCCAG 552
DB 870 CGTATCTGCTGACATTTGGCGCTGCCCATGAACCGTAAGCTGTATCCGCTGGATAGACAG 929
OY 553 ACGTCAGAGATGCAAGCTTGAAGCTCATTCATCTGACAGCCCTCTCCATCTCTGCA 612
DB 930 ATCTGCTCTACATGAGATGCGCAGCT----- 954
OY 613 CTTTCAGTTGGCTACACCATGAAAGACCTCGTTTGTAGTGCGTGAAGATGCTCCCTGCT 672
DB 955 -----ATGGCTGGAGACCACACAGCACTTGGCTCTTCCCTGTGGAAGAGAGGGGATC---CC 1004
OY 673 GTCAAGTGGCTGAGAGGGCTGACTCTGCCCCAGATTATCTTCCGGATGAGAAGGATGTA 732
DB 1005 GTACAGAGGTGAAGAACTTACACCTACCTCGCTTCACTGAGAGAACTTCTGACT--- 1061
OY 733 GGCCTGTCTACCAAGCACTACACACAGGAAATTCACCTCGATCGAGTGAAGTTTAC 792
DB 1062 GATTACTGTAACTAAGAAACCAACACCGGTGAATACAGTGCCTCAAGTGCATCTACTA 1121
OY 793 CTGGAACGCGAGATGGGCTACTATCTGATGATGATGATGCCAGGCTACTGCTGCTC 852
DB 1122 TTCAGCGCAAGATTTCTCATTTACTTATACAAATTTATATACCATGCTGTATGTGTC 1181
OY 853 ATCTGCTCTGGGTCTCTCTCTGATCAACATGATGCTGCCCTGCCGCTGTGGGCTG 912
DB 1182 ATTGATCATGGATATCATCTTGGCTGTGATCAAGAGAGATACGGGCGGATGTCACTG 1241
OY 913 GGCAATCACCACCGTCTACCATGACCAACAGAGCTTGGCTCCCGGCTCTTTGGCT 972
DB 1242 GGTGTACACACCGCTGACCATGGCCACCAAGAGTGGGCTAAGGCTCCCTGCCCTG 1301
OY 973 AAGGTCTCTAGTGAAGCAATGCAATCTGATGGCTGTGTGCTCTTTGCTTC 1032
DB 1302 CCCGTTTCTTATAGCAAGCCATGATGTGTGAGACAGGCGTGTCTACAGTTGCTGTC 1361
OY 1033 GCTGCTTGTGAGATGATGCTGCAATAATTTGTTCTCG 1073
DB 1362 GGGGCCCTGCTGAGTTCGCCCTGGTGAATATGATCCCG 1402

```

RESULT 6

Sequence 5, Application PC/TUS9606035

GENERAL INFORMATION:

APPLICANT: Cully, Doris F.

APPLICANT: Arena, Joseph P.

APPLICANT: Pares, Philip S.

APPLICANT: Liu, Ken K.

TITLE OF INVENTION: DNA ENCODING GLUTAMATE GATED CHLORIDE

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: Jody M. Glesser

STREET: 126 East Lincoln Avenue - P.O. BOX 2000-0907

CITY: Rahway

STATE: New Jersey

COUNTRY: US

ZIP: 07065-0907

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US96/06035

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Glesser, Jody M.

REGISTRATION NUMBER: 32,838

REFERENCE/DOCKET NUMBER: 19264 PCT

TELECOMMUNICATION INFORMATION:

```

: TELEPHONE: (908) 594-3046
: TELEFAX: (908) 594-4720
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3958 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: PCT-US96-06035-5.

```

Query Match

Best Local Similarity 53.7%; Pred. No. 1.4e-42;

Matches 473; Conservative 0; Mismatches 360; Indels 48; Gaps 3;

```

OY 193 AATTTAAAGGCCACCCCGTGAAGCTGACCTGCAACACTTTCATCAACAGTTTCAAGCTCC 252
DB 570 AATGGACAGATGTGTCGCCCATAGTCAATCAATCTATTCTTCCGAGATTAATGACG 629
OY 253 GTCAACAGACACCAATGACTACCGGGTGAATGCTTCTGGGCAACAGCTGAATGAC 312
DB 630 ATTAGTGTATTTAAATGAGTACAGTGTGCAAGTTAACCTTCCGGAACAGTGAAGCAT 689
OY 313 CCAGGCTGTCTAACCGAATATCTGATGATGATGATGATGATGATGATGATGATGATG 372
DB 690 GAAGGCTCAAGATTCAGATATCCAGGGTGGCCCTAAGTATGATGATGATGATGATGATG 749
OY 373 GACTCTATCTGGAAGCCAGACCTCTTCTTCTGTAATGAGAAAGGGGCCAATCTTCATG 432
DB 750 AAGCGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 809
OY 433 GTGACACAGCAACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 492
DB 810 ATCATCATGCCCCAATGTGTATATTTGCAATCTTCCCAACGATCTGTCTATATAGTATA 869
OY 493 AAGGTACCTCTATTTTGTCTCTGCTGATGATGATGATGATGATGATGATGATGATGAT 552
DB 870 CGTATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 929
OY 553 ACGTCAGAGATGCAAGCTTGAAGCAATGCAATGATGATGATGATGATGATGATGATGAT 612
DB 930 ATCTGCTACTAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 954
OY 613 CTTTCAGTTGGCTACACCATGAAAGACCTCGTTTGTGAGTGGCTGAAGATGCTCTGCT 672
DB 955 -----ATGGCTGGAGACCACCAAGCACTTGGCTCTTCCCTGTGGAAGAGGGGATC---CC 1004
OY 673 GTCAAGTGGCTGAGAGGGCTGACTCTGCCCCAGATTATCTTCCGGATGAGAAGGATGTA 732
DB 1005 GTACAGGTGTGAAGAACTTACACCTACCTGCTTCACTGAGAGAAAGTTCTGACT--- 1061
OY 733 GGCCTGTCTACCAAGCACTACACACAGGAAATTCACCTCGATCGAGTGAAGTTTAC 792
DB 1062 GATTACTGTAACTAAGAAACCAACACCGGTGAATACAGTGCCTCAAGTGCATCTACTA 1121
OY 793 CTGGAACGCGAGATGGGCTACTATCTGATGATGATGATGATGCCAGGCTACTGCTGCTC 852
DB 1122 TTCAGCGCAAGATTTCTCATTTACTTATACAAATTTATATACCATGCTGTATGTGTC 1181
OY 853 ATCTGCTCTGGGTCTCTCTCTGATCAACATGATGATGATGATGATGATGATGATGATGAT 912
DB 1182 ATTGATCATGGATATCATCTTGGCTGTGATCAAGAGAGATACGGGCGGATGTCACTG 1241
OY 913 GGCATCACCACCGTCTACCATGACCAACAGAGCTTGGCTCCCGGCTCTTTGCTCT 972
DB 1242 GGTGTACACACCGCTGACCATGGCCACCAAGAGTGGGCTAAGGCTCCCTGCCCTG 1301
OY 973 AAGGTCTCTAGTGAAGCAATGCAATCTGATGATGATGATGATGATGATGATGATGATGAT 1032
DB 1302 CCCGTTTCTTATAGCAAGCCATGATGTGTGAGACAGGCGTGTCTACAGTTGCTGTC 1361
OY 1033 GCTGCTTGTGAGATGATGCTGCAATAATTTGTTCTCG 1073

```





Db	573	-----TGGTTGAGCCAAACGATCTGGT	596
OY	645	GTTTGAATGGCTGGAAGATGCTCTGCTGTCACAGTGGCTGAGGGGCTGACTCTGCCCA	704
Db	597	TTTTTTTGTGAAGGAAAGGTGACCCG---GTGCAGGTGTCAAGATCTACATCTGCCAG	653
OY	705	GTTATATCTCGGGATGAGAAGATCTTAGGCTGTTGTACCAAGCATTAACAACACAGGAA	764
Db	654	GTTTACGTT---GGAGAAGTTCTTGACGGAATTATTGTACAGCAAAACCAATACCGGTGA	710
OY	765	ATTACCTGCATCGAGGTAAAGTTTACCTGGAAGCGCAGATGGGCTACTATCTGATTC	824
Db	711	ATACAGTTGGCTGAAGGTGCACGTCCTTTAAACGAGAATTCTCGTACTACCTGATCCA	770
OY	825	GATGTACATCCCCAGCCTACTCATCGCATCTCTGCTGGGTCTCTTTGGAATCAAT	884
Db	771	GATCTACATTCCTTTGTCATGTTGGATGATGTTCCCTGGGTGTCGCTTGTTGGAGCA	830
OY	885	GGATGTCGCCCTGCGCGGTGGGCGTCGGGCATCACACGCTGTCACACATGACACCCA	944
Db	831	GGGAGCGGTTCCGGCCAGAGATATACGTAGGGGTGATACCATCTCTCCACACATGACGCCACCA	890
OY	945	GAGCTCTGGCTCCCGGGGCTCTTGTCCCTAAAGGTGCTACGTGAAGCATGACATCTG	1004
Db	891	GACGTGGGCATTAACGCTCCCTCCGCGGCAGGTGCTCTACACAAAGCATGACGCTCG	950
OY	1005	GATGGCTGTGTGTGCTCTTTTGTGTTCGCTGCTTGGATGATGCTGCCATTAATTT	1064
Db	951	GACCGGAGTCTGCCTCACGTTCTGTTCTGGGGCTTGTCTGGAATTCGCCCTGCTCACTA	1010
OY	1065	TGTTTC 1070	
Db	1011	CGCCTC 1016	
RESULT 9			
US-09-130-339-3			
: Sequence 3, Application US/09130339			
: Patent No. 6358701			
GENERAL INFORMATION:			
: APPLICANT: Cully, Doris F.			
: APPLICANT: Pareess, Philip S.			
: APPLICANT: Warmke, Jeffrey W.			
: APPLICANT: Etter, Adrian			
: APPLICANT: Cohen, Charles J.			
: APPLICANT: Brochu, Richard M.			
: TITLE OF INVENTION: DNA MOLECULES ENCODING CTENOCEPHALIDES			
: TITLE OF INVENTION: FELIS GLUTAMATE GATED CHLORIDE CHANNELS			
: FILE REFERENCE: 20029			
: CURRENT APPLICATION NUMBER: US/09/130.339			
: CURRENT FILING DATE: 1998-08-06			
: NUMBER OF SEQ ID NOS: 8			
: SOFTWARE: FastSeq for Windows Version 3.0			
: SEQ ID NO 3			
: LENGTH: 1297			
: TYPE: DNA			
: ORGANISM: ctenocephalides felis			
US-09-130-339-3			

	Query Match	9.8%	Score 161.2;	DB 4;	Length 1297;
	Best Local Similarity	53.9%	Pred. 1.4e-37;		
	Matches	433;	Conservative	0;	Mismatches 322; Indels 48; Gaps 3;
OY	267	AATGACTACCGGGTGAATGTCTTCCTTGCGGCACAGTAGAATGCCAGCCTGTCTTA	326		
Db	190	ACTGGAATACAGTGTCCAGTTAACTTTCAGGAACAATGCCAGGTGAGAATTGAANAAT	249		
OY	327	CCGAGAAATATCCTATGACTCTCTGGACCTGCATCCCTCAATGCGACTATCTGGAA	386		
Ddb	250	TAAAGACTTGGAGAGCTGTTAAAATAATCTTAACACTAACGACAGCAATCGTATGGAT	309		
OY	387	GCCAGACCTCTTCTTGGTAAATGAGAAAGGGCCAATCTCATGATGAGTGACCAGGACAA	446		

D	b		310	GCCCCGATTGGTCTTTGGCAGTAAGAAAAGAGGGCCACTTTCACATCATCATCGCCGA	369
O	y		447	CAAGTACTGGGCATCTTCAAGAATGGGAATGTGCTGTACAGATCAGAGCTACCCCTAT	506
D	b		370	CGTCTACATTCGTATTTTTCCTTACGGTTCCGTACTATAACACATCAGATATVCCCTTAC	429
O	y		507	TTTTGCTGCCGTGATNGACCTCAAGAACTTCCCCTATGGACATCCAGAGTCGACGATGA	566
D	b		430	TTTBGGCGTGCTCTATGAATCTGMAAAGCTGATTCGCTGATAGGACAGTGTGCTCTCG	489
O	y		567	GCTTGAGAGCTCATCCATCATCTCTGCAGCCCTCTGCCATCTGTGCATTTACGTTGGCTA	626
D	b		490	GATGGCCAGTTA-----TGGTGTG	507
O	y		627	CACCATGAAGAACCCGTGTTTGATGGGCTGGSAAGATGCTCCTGCTGTCACAGTGGCTGA	686
D	b		508	GACCACAACACATCTGGTGTTTTGTGGAAGSAGGTGACCCG--GTGCAGGTGTCA	564
O	y		687	GGGGCTGACCTCGCCCCAGTTTATCTTCGGGATGAGAAGATCTTAGGCTGTTHAACAA	746
D	b		565	GAATCTACATCTGCCCCAGGTTTACGTT--GGAGAAGTCTTGTACGAGATTAATGTAAACG	621
O	y		747	GCACTACACACAGGGAATTCACCTGCATCGAAGSTAAAGTTTACACTGGAACGGCAGAT	806
D	b		622	CAAAACCAATACCCGGTGAATACAGTTGGCTGGAAGGTGCACCTGCTTTTAAACGAGAGTT	681
O	y		807	GGGCTACTATCTGATTCAGATGTACATCCCAAGCCCTACTCATCATCCGTGCGGGT	866
D	b		682	CTCGTACTACCTGATCCAGATCTACATTCCTTGTGGCATGTTGGTATGCTTTCCGGGT	741
O	y		867	CTCCTTCTGSGATCAACATGATGCTGCCCCCTCCCGTGTGGGCTGTGGCATCACACCGT	926
D	b		742	GTCGTCTGTGGTTGACCAAGGAGCCGTTTCCGGCCAGAGATACATCGGGGTGACACCTCT	801
O	y		927	GCTCAGCATGACACACCAGAGCTCTGGGCTCCGGGACCTTTTGCCCTAAGGTGCTACGCT	986
D	b		802	CCTCAGCATGGGCAACCCAGAGCTGGGGCATAAACGCTCTCCGTCCGGCCAGTGTCTACAC	861
O	y		987	GAAAGCAATCGACATCTGGATGCTGTGTGCTCTTTGTGTTCGCTGCTGCTGCA	1044
D	b		862	AAAAGCCATCGACGCTGGACGGAGTCTGCCCTCACGTCGTCCTTCGGGGCTTTGCTCGA	921
O	y		1047	GTAAGCTGCCATAAATTTTGTGTT	1070
D	b		922	ATTCGCCCTCGTCACTACGCTTC	945

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

ATTORNEY/AGENT INFORMATION:  
NAME: Digilio, Frank S  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 9732  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1657 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 115..1581  
US-08-554-659-3

Query Match 9.6%; Score 157.8; DB 1; Length 1657;  
Best Local Similarity 51.6%; Pred. No. 1.6e-36;

Matches 474; Conservative 0; Mismatches 397; Indels 48; Gaps 3;

```

OY 137 ATTCTAGACAACTTATGGGCGAATCTGGATATGATGCCAGATTGCGCCCAATT 196
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 239 ATATCTCAGCCATTTTGGATTGCTAAGTAAAGTAAAGAGAGAGAGGCGCAACT 298
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 197 TTAAGGCGCCCGGAGACGTAAGTACCTGACATCTTCATCAAGTTTCACTCCGTCA 256
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 299 ATGGAGAGCCGAGGATGAGTGGAGATCAACATGATGCTCTCCATCACTCTCTTAT 358
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 257 CCAAGACCAATGAGTACCGGCGAATGCTTCTTGGCGGACAGATGAGACCCAC 316
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 359 CTGAGTGAATGATTTTACCCCTGGATTTCTACTGACAAATTTTGGACAGACCCCA 418
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 317 GCCTGCTCCACGAAATATCC---TGATGACTCTCTGACCTCCATCCCTCAATGCTG 373
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 419 GCTTGTCTTCAAAAAGACGAGGAGTGGAGACTGTCTCCCTCGGCTCGAATTTATTA 478
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 374 ACTTATCTGAGAGCAGACTCTTCTTGTATGATGAAAGGGGCACTTCCATGAGG 433
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 479 GAAACATATGGGTACCCGACACTTCTTGTATGAGAAAGATGATTTTCCACATAG 538
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 434 TGAACACGAGCAAGATTACTGCGCATCTTCAAGATGGAGTGTCTGTACAGCATCA 493
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 539 CTACTACAGCAAGATTCATACGATTCATCATCTGATCTGATCTATTACTAGAGATA 598
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 494 GGCTGACCTCATTTTGTCTGATGAGACCTCAAGAACTTCCCATGAGACATCCAGA 553
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 599 GACTGACTATACCCGCTTCTTGTCCGATGATTTGCAAGTATTTTCCGATGACCGTCAAT 658
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 554 CCTGACAGTACAGCTTGAAGCTCATCTGACAGCCCTGACATCTGTCTAC 613
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 659 TATGCAATTTGAATCGAAAGTT----- 682
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 614 TTTCAAGTTGGCTACACCATGAAACCTGCTTGTGATGGCTGGAAGATGCTCTGCTG 673
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 683 -----TTGGCTTACACATCGCGGACATCGATACAAAGATGAGGGGCCCAACTAG 736
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 674 TCCAGTGGCTAGGGGCTGACTTGGCCCACTTATCTTGGGGATGAGAGATCTAG 733
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 737 TGGGTGTGCGAGCAAGTGTCTTGGCCCAATTTCAAGTGTGGGCCACCGGACGGGG 796
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 734 GCTGTGTACAGCACTACACAGAGGAATTCACCTGCTGACGATGAGTAAAGTTTCCAC 793
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 797 CCATGAGAGTTTCTC---TTAGACAGAGAACTACTCTGCTGTGCGATGTGAATTTCAAT 853
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 794 TGAACGCGAGATGGCTACTATCTGATTCAGATGATATCCCAAGCCCTACTCATCTGTA 853
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 854 TTGTAGGCTCGATGGATCTATTTAATTCAGATTTATATTCGCTGCTGCTAATTTGTA 913
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 854 TCTGTCTCTGGGTCTCTTCTGTGATCAACATGATGATGCTCCCTGCGCTGTGGGCTG 913
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

DB 914 TATATCTTGGGTATCATTTTGTGAAATGCAATGCGACACCTGCAAGGGTAGACATAG 973
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 914 GCATCACCCCTGCTCACCACATGACACCCAGACGCTGTCCCGGCGCTTTGGCTTA 973
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 974 GTGTCAACAATGATTTAGACGACGCTCATGTCTGTCCAGAAATGGGCTTGCCCA 1033
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 974 AGTGTCTTACGTGAAGCAATCGACATCTGATGATGCTGTGCTCTTTGTGTTG 1033
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1034 AATCTCATATGTCAAGTCAATCGATGATCTATCTGGAACCTTTTGTGATGCTTCA 1093
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1034 CTGCTTGTGATGATGC 1052
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1094 CCAGTTTACTAGAAATATGC 1112
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

## RESULT 11

US-09-002-361-1

Sequence 1, Application US/09002361  
Patent No. 6329516

GENERAL INFORMATION:

APPLICANT: Halling, Blak  
TITLE OF INVENTION: Lepidopteran GABA-gated Chloride

NUMBER OF SEQUENCES: 43

CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Dechert Price & Rhoads

STREET: 997 Lenox Drive, Building 3, Suite 210

CITY: Lawrenceville

STATE: ND  
COUNTRY: USA

ZIP: 08543

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/002,361  
FILING DATE:

CLASSIFICATION:  
PRIORITY APPLICATION DATA:

APPLICATION NUMBER:  
FILING DATE:

ATTORNEY/AGENT INFORMATION:  
NAME: Bloom, Allen

REGISTRATION NUMBER: 29,135  
REFERENCE/DOCKET NUMBER:

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 609-520-3214

TELEFAX: 609-520-3259  
TELEX:

INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:

LENGTH: 1844 base pairs  
TYPE: nucleic acid

STRANDEDNESS: single  
TOPOLOGY: linear

FEATURE:  
NAME/KEY: Coding Sequence

LOCATION: 104...1591  
OTHER INFORMATION:

US-09-002-361-1

Query Match 9.6%; Score 157.8; DB 4; Length 1844;

Best Local Similarity 51.6%; Pred. No. 1.7e-36;  
Matches 474; Conservative 0; Mismatches 397; Indels 48; Gaps 3;

```

OY 137 ATTCTAGACAACTTATGGGCGAATCTGGATATGATGCCAGATTGCGCCCAATT 196
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 228 ATATCTCAGCCATTTTGGATTGCTAAGTAAAGTAAAGAGAGAGGCGCAACT 287
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 197 TTAAGGCGCCCGGAGACGTAAGTACCTGACATCTTCAATCAACAGTTTCACTCCGTCA 256
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Db 288 ATGAGAGCCCGCATGATGTGGAGTCACCATGATGCTCTCCATCAGCTCTTAT 347  
Qy 257 CCAAGACCAATGAGTACCGGGTGAATGCTCTTCCGCAACAGGAGATACCCAC 316  
Db 348 CTGAGTGAATAATGATTTACCCCTGGATTTCTACTTACGACAAATTTTGGACACACCCA 407  
Qy 317 GCTGTCTTACCGAATATCC--TGATGACTCTTGAGACTTGATCCCTCCATGCTGG 373  
Db 408 GCGTGGCTTACAAAAAAGACGGGTGGAGACTCTGCTGCGCTCGGAATTTATTA 467  
Qy 374 ACTGATCTGGAACCCAGACCTCTTCTGCTAAATGAGAAAGGGCCAACTTCATGAG 433  
Db 468 GAAACATATGGGTACCCGACACCTTCTTTGTTAAAGAAAAACAGCTTATTTCCACATAG 527  
Qy 434 TGACACGAGCAACAAGTACTGGGCATCTTCAAGAAATGGAAATGCTGTACACATCA 493  
Db 528 CTACTACAGCAAGAAATTCATACGCAATTCATCTTGATCTAATTAAGAGATATA 587  
Qy 494 GCGTGAACCTCTATTTCTCTGCTGATGACCTCAAGAACTTCCCATGACATCCAGA 553  
Db 588 GACTGACATACCGCTTCTTGTCCGATGATTTGGACATATTTCCGATGACCGTCAAT 647  
Qy 554 CGTGCAGATGACCTGAGAGCTCATCTGACAGCCCTCTGCAATCTCTGTCAC 613  
Db 648 TATCCAATATGAAATGAAAGTT----- 671  
Qy 614 TTTGAGTTGCTACCATGAAACCTCTGTTGATGAGTGGCTGGAAGATCTCTGCTG 673  
Db 672 -----TTGGCTACACATGCGGAGATCCGATACAACTGAAATGAGGGCCCAACTGAG 725  
Qy 674 TCCAAGTGGCTGAGGGGCTGACTGCCCCAGTTTATCTTGGAGATGAGAAGATCTAG 733  
Db 726 TGGGTGTGCAAGGAGTGTCTTCCGCAATTCAGAGTCTGAGCCACCGGCAAGCGGG 785  
Qy 734 GCTTTTGTACCAAGCACTACAAACAGGAAATTCACCTGATGAGGTAAGTTTCACC 793  
Db 786 CCATGAGATATTC---TTACGACAGGAACACTCTGCTGAGATGAAATTCAT 842  
Qy 794 TGGAAAGGCAATGGGCTACTATCTGATTCAGATGATATCCCAAGCTTACTCATGCTCA 853  
Db 843 TTTGAGCTGATGGATGATATTTAATTCAGATTTAATTCCTGCGCTTAATTTGCA 902  
Qy 854 TCCGTCTGCTGCTCTCTTGTGATCAACATGATGCTGCCCTGCGCGCTGAGGCTCGG 913  
Db 903 TTAATCTTGGGTATCATTTTGGTTGAATGCAAAATGCGACACTGCAAGGATACACTAG 962  
Qy 914 GCATCACCACCGTCTCACCATGACCAACCCAGAGCTGCTGCCGGGCTCTTTGCTTA 973  
Db 963 GTGTACAACTGTATTGACATGACAGCGTCTGCTCCAGCAATGCGGCTGCGCCA 1022  
Qy 974 AGGTGCTCTACGTAAGGCAATGCAATCTGATGCGCTGTCTGCTCTTTGTGCTG 1033  
Db 1023 AGATCTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1082  
Qy 1034 CTGCTTCTGCTGAGTATGC 1052  
Db 1083 CCAGTTTACTGAAATATGC 1101

RESULT 12  
US-08-417-330A-19  
: Sequence 19, Application US/08417330A  
: Patent No. 5719057  
: GENERAL INFORMATION:  
: APPLICANT: HADINHAM, KAREN  
: APPLICANT: LE BOURDELLES, BEATRICE  
: APPLICANT: WHITING, PAUL  
: APPLICANT: WINGROVE, PETER  
: TITLE OF INVENTION: STABLY TRANSDUCED CELL LINE EXPRESSING  
: TITLE OF INVENTION: GABA-A RECEPTOR AND NOVEL CLONED  
: NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:  
ADDRESSEE: J. MARK HAND - MERCK & CO., INC.  
STREET: 126 EAST LINCOLN AVENUE - P.O. BOX 2000  
CITY: RAHWAY  
STATE: NJ  
COUNTRY: US  
ZIP: 07065-0900  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/417,330A  
FILING DATE: 05-APR-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: HAND, MARK  
REGISTRATION NUMBER: 36,545  
REFERENCE/DOCKET NUMBER: T11091A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 908-594-3905  
TELEFAX: 908-594-4720  
TELEX:  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1866 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: Coding Sequence  
LOCATION: 225...1646  
OTHER INFORMATION:  
US-08-417-330A-19

Query Match 9.5%; Score 156.4; DB 1; Length 1866;  
Best Local Similarity 52.0%; Pred. No. 4.3e-36;  
Matches 460; Conservative 0; Mismatches 376; Indels 48; Gaps 3:

Qy 169 GGAATATGATGCCAGATTCGGCCCAATTTTAAAGCCCGGAGACGTACCTGCAC 228  
Db 360 GCGTATGACATTCGCTGAGACAGATTTTGGAGGTCCCGCTGGGGATGAGAC 419  
Qy 229 ATCTTCAATCAACAGTTTCAGCTCCGTCACAAAGACACATGAGTACCGGGTGAATGTC 288  
Db 420 ATTGACATTTGGCAGCATGATGATGATGATGATGATGATGATGATGATGATGATG 479  
Qy 289 TTTCTGGGCAACAGTGAATGACCCAGCGCTGTCTTACCGAGAAATATCTGATGACTCT 348  
Db 480 TACTTTCAACAAGCTGAGAGATTAAGAGCTGTCTTAATTAATTAATTAATTA 537  
Qy 349 CTGACCTGATTCCTCTCATGCTGAGCTCTTCTGGAAGCCAGACCTCTTTGCTAAT 408  
Db 538 -TGACTTGGCAACAGAGTGGACAGCGCTGAGGCTGATTAATTAATTAATTAATTA 596  
Qy 409 GAGAAAGGGCCCAACTTCATGATGAGTGCACAGCAACAAGTACTGCGCATCTTCAAG 468  
Db 597 GATAAAGATGATTTGTGTCAGAGAGTGTCTTAAAGAACCGCATGATTCGCTGATCT 656  
Qy 469 AATGGAAATGCTGTACAGCATGAGCTGACCTCATTTTGTCTGCTGATGAGACTC 528  
Db 657 GATGGACCGCTCTTATGAGACTCAGAAATCAACAACAGCTGCTGATGAGACTCA 716  
Qy 529 AAGAACTTCCCATGAGCATGCAACGTCAGCATGAGCTTGAAGAGCTCATCATATC 588  
Db 717 AGGAGTATCCATGAGTGAACAAACTGCACTTGGAAATTTGAGAGCTA----- 766  
Qy 589 TGCAGCCCTTGCACATCTCTGCACTTTCAGTTGCTTACACAGAAAGACCTGCTGTT 648





FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Bloom, Allen  
REGISTRATION NUMBER: 29,135  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 609-520-3214  
TELEFAX: 609-520-3259  
TELEX:  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 669 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: Coding Sequence  
LOCATION: 2..667  
OTHER INFORMATION:  
US-09-002-361-7

Query Match 8.9%; Score 146.6; DB 4; Length 669;  
Best Local Similarity 54.6%; Pred. No. 1.9e-33;  
Matches 367; Conservative 0; Mismatches 260; Indels 45; Gaps 2;

OY 399 CTTTGTATGAGAAAGGGGCACTTCATGAGGTGACACGAGACAACTACTGCG 458  
DB 1 CTTTGTATGAGAAAGGGGCACTTCATGAGGTGACACGAGACAACTACTGCG 60  
OY 459 CATCTTCAAGATGGGAATGCTGTACACATCAAGCTGACCTCATTTTGTCTGCT 518  
DB 61 CATCCACCACTGGGCTCATCAGCGTACATAGGCTCACCATCAGGCTCTCTGCCC 120  
OY 519 GATGACCTCAAGAACTTCCCATGACATGACATGACATGACATGACATGACATG 578  
DB 121 CATGAACCTGACGACTTCCCATGATGACATGACATGACATGACATGACATGAC 179  
OY 579 ATCCATACCTGACAGCCCTCTCCATCTGTCTACTTGTGCTACATGACATGAC 638  
DB 180 -----TCGGCTACACCATGCGGA 198  
OY 639 CCTCGTTTGAAGTGGCTGAGATGCTCTCTGCTGCAAGTGGCTGAGGGCTGACTCT 698  
DB 199 CATCCGGTACAAATGAGACGAGGGGNCACACTCGGTGGCGTTTCAAGCAAGTGTCT 258  
OY 699 GCCCAGTTTATCTTGGGGATGAGAGGATCTAGGCTGTGTACCAAGCACTACAAAC 758  
DB 259 ACCGAGTTCAAGTGTGGGCTGACATGTCACGTCGATGAATAATATGCT--CACAC 315  
OY 759 AGGGAATTCACCTGATGAGTAACTTACCTGAGACGAGATGGCTACTATCT 818  
DB 316 AGGAACCTACTCCGCGGTGGCGTGGAGTCAAGTGTGCTGATGGCTACTACTCT 375  
OY 819 GATTGATGATACATCCCGAGCTACTCATCTGCTGCTGCTGCTGCTGCTGCTGCTG 878  
DB 376 GATCCAGATCTACATACCAATCGCGGCTCATGCTCAATATATGCTGCTGCTGCT 435  
OY 879 CAACATGATGCTGCGCCCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 938  
DB 436 GAACCGCAAGCGAGCGCGCGCGCGGTGAGCTGGGCGTCAACCGTCTCACCATGAC 495  
OY 939 CACCCAGACCTGCTCCCGGCTCTTGGCTAAAGTGTGCTGCTGCTGCTGCTGCTG 998  
DB 496 CACGCTCATGCTTCTCAATATGCGCGCTGCGGAGATCTGTAGCTTAAGTCCATGCA 555  
OY 999 CATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1058  
DB 556 TGTGTACCTGCGGACCTGCTGCTATATGCTTACACCAATCTGCTAGAGTACGCGAG 615

OY 1059 AAATTTGTTTC 1070  
DB 616 GCGTATATGTC 627

Search completed: June 30, 2003, 23:02:26  
Job time : 116 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OW nucleic - nucleic search, using sw model

Run on: June 30, 2003, 21:06:15 ; Search time 261 Seconds  
(without alignments)  
9327.568 Million cell updates/sec

Title: US-10-075-846-3  
Perfect score: 1640  
Sequence: 1 atgcaactctgtctctgc.....taaaaaaaaaaaaaaaaa 1640

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1055720 seqs, 742224136 residues

Total number of hits satisfying chosen parameters: 2111440

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_NA:\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCR\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCIT05\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1640	100.0	1640	US-10-075-846-3	Sequence 3, Appl1
2	1524.8	93.0	2565	US-10-075-846-1	Sequence 1, Appl1
3	807.8	49.3	993	US-10-075-846-9	Sequence 9, Appl1
4	480.6	29.3	679	US-10-075-846-5	Sequence 5, Appl1
5	219	13.4	492	US-09-918-995-15843	Sequence 15843, A
6	193	11.8	1555	US-10-211-673-11	Sequence 11, Appl1
7	187.6	11.4	1197	US-10-239-420-11	Sequence 10, Appl1
8	181.6	11.1	4621	US-09-969-844-13	Sequence 13, Appl1
9	181.6	11.1	1609	US-09-969-844-11	Sequence 11, Appl1
10	174.2	10.6	1640	US-09-969-844-12	Sequence 12, Appl1
11	171.8	10.5	3442	US-10-239-420-3	Sequence 3, Appl1
12	168.6	10.3	1467	US-09-815-925-10	Sequence 10, Appl1
13	167	10.2	3598	US-10-239-420-1	Sequence 1, Appl1
14	162.6	9.9	1458	US-09-808-483-11	Sequence 11, Appl1
15	162.6	9.9	1608	US-09-808-483-9	Sequence 9, Appl1
16	155	9.5	480	US-09-864-761-11304	Sequence 11304, A
17	153.4	9.4	569	US-09-899-495-19	Sequence 19, Appl1
18	152	9.3	1150	US-10-239-420-13	Sequence 13, Appl1
19	151.6	9.2	455	US-09-864-761-5998	Sequence 5998, Ap

20	150.8	9.2	2194	US-10-239-420-4	Sequence 4, Appl1
21	143.2	8.7	1997	US-10-037-270-256	Sequence 256, App
22	142.2	8.7	4077	US-10-239-420-6	Sequence 6, Appl1
23	138.8	8.5	1192	US-09-808-483-5	Sequence 5, Appl1
24	138.8	8.5	1269	US-09-808-483-3	Sequence 3, Appl1
25	138.8	8.5	1272	US-09-808-483-7	Sequence 7, Appl1
26	137	8.4	1323	US-09-510-662A-38	Sequence 38, Appl1
27	137	8.4	1323	US-09-778-320-38	Sequence 38, Appl1
28	137	8.4	1323	US-09-910-689-38	Sequence 38, Appl1
29	137	8.4	1323	US-10-010-742-38	Sequence 38, Appl1
30	137	8.4	2593	US-10-106-698-2031	Sequence 2031, Ap
31	137	8.4	3282	US-09-954-531-1034	Sequence 1034, Ap
32	137	8.4	3282	US-09-825-301-75	Sequence 75, Appl1
33	136.2	8.3	1038	US-09-765-069-9	Sequence 9, Appl1
34	136.2	8.3	1179	US-09-765-069-3	Sequence 3, Appl1
35	136.2	8.3	1263	US-09-765-069-7	Sequence 7, Appl1
36	136.2	8.3	1404	US-09-765-069-1	Sequence 1, Appl1
37	136.2	8.3	1422	US-09-742-311-1	Sequence 1, Appl1
38	136.2	8.3	1987	US-10-037-270-331	Sequence 331, App
39	135.2	8.2	1052	US-09-969-844-8	Sequence 8, Appl1
40	130.6	8.0	1693	US-10-037-270-131	Sequence 131, App
41	130.6	8.0	1866	US-09-964-824A-96	Sequence 96, Appl1
42	128.6	7.8	184	US-09-864-761-28002	Sequence 28002, A
43	127	7.7	1398	US-09-893-321-1	Sequence 1, Appl1
44	127	7.7	1568	US-09-893-321-5	Sequence 5, Appl1
45	125.8	7.7	2819	US-09-818-657-1	Sequence 1, Appl1

#### ALIGNMENTS

```
RESULT 1
US-10-075-846-3
; Sequence 3, Application US/10075846
; Publication No. US20030032608A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING A NOVEL GLYCINE RECEPTOR ALPHA SUBUNIT
; TITLE OF INVENTION: IN THE GASTROINTESTINAL TRACT, HGRA4, and SPLICED VARIANT THERE
; FILE REFERENCE: D0079 NP
; CURRENT APPLICATION NUMBER: US/10/075.846
; CURRENT FILING DATE: 2002-02-13
; PRIOR APPLICATION NUMBER: US 60/269,535
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 1640
; TYPE: DNA
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1293)
US-10-075-846-3

Query Match      100.0%  Score 1640:  DB 9:  Length 1640:
Best Local Similarity 100.0%:  Pred. No. 0:
Matches 1640:  Conservative 0:  Mismatches 0:  Indels 0:  Gaps 0:
```

QY 1 ATGACAACTCTTGTTCGCAACCTCTTCCTTCCTTCGCAACCTCGCAGGCGAG 60  
|||||  
DB 1 ATGACAACTCTTGTTCGCAACCTCTTCCTTCCTTCGCAACCTCGCAGGCGAG 60  
|||||

QY 61 GTTCTCTCAGGAGGCTTGCAAAAGAGCAATCTGCAACCAAGGCGCGAG 120  
|||||  
DB 61 GTTCTCTCAGGAGGCTTGCAAAAGAGCAATCTGCAACCAAGGCGCGAG 120  
|||||

QY 121 CCCATGTCCTCCCTGATTTCTAGCAAACTTATGGGCGCAATCTGATATGATGCC 180  
|||||  
DB 121 CCCATGTCCTCCCTGATTTCTAGCAAACTTATGGGCGCAATCTGATATGATGCC 180  
|||||

QY 181 AGGATTCGGCCCAATTTAAAGCCCGCTGACGTCGACATCTTCATCAAC 240  
|||||



Db 181 AGGATTGGGCCCAATTTTAAAGGCCACCGGTGACGTGACATCTTTCATCAAC 240  
OY 241 AGTTAGAGTCCCTGCAACAAGACACATGAGTACCGGGTGAATGCTTCTGGGCA 300  
Db 241 AGTTAGAGTCCCTGCAACAAGACACATGAGTACCGGGTGAATGCTTCTGGGCA 300  
OY 301 CAGTGAATGACCCACGCGCTGCTCAACGGAATATCTGATGACTCTGACCTGCAT 360  
Db 301 CAGTGAATGACCCACGCGCTGCTCAACGGAATATCTGATGACTCTGACCTGCAT 360  
OY 361 CCGTCATGCTGAGCTATCTGGAAGCCAGACCTTCTTCTGATGAGAAAGGGCC 420  
Db 361 CCGTCATGCTGAGCTATCTGGAAGCCAGACCTTCTTCTGATGAGAAAGGGCC 420  
OY 421 AACTTCATGAGTGAACAAGCAAGTACTGCGCATCTTCAAGAAATGGGAATG 480  
Db 421 AACTTCATGAGTGAACAAGCAAGTACTGCGCATCTTCAAGAAATGGGAATG 480  
OY 481 CTGTACAGCATCAGGCTGACCTCATTTTGTCTGCTGATGAGACTCAAGAACTTCCC 540  
Db 481 CTGTACAGCATCAGGCTGACCTCATTTTGTCTGCTGATGAGACTCAAGAACTTCCC 540  
OY 541 ATGACATCCAGCGTGAAGGACGATGAGCTGAGAGCTCATCTGACCTGACCCCTG 600  
Db 541 ATGACATCCAGCGTGAAGGACGATGAGCTGAGAGCTCATCTGACCTGACCCCTG 600  
OY 601 CCATCTCTGCTACTTTCAGTTGGCTACACCATGAAGACCTGCTTTGATGGCTGAA 660  
Db 601 CCATCTCTGCTACTTTCAGTTGGCTACACCATGAAGACCTGCTTTGATGGCTGAA 660  
OY 661 GATGCTCTGCTCTCCAGTGGCTGAGGGCTGACTCTGCCCCAGTTTATCTTGGGGAT 720  
Db 661 GATGCTCTGCTCTCCAGTGGCTGAGGGCTGACTCTGCCCCAGTTTATCTTGGGGAT 720  
OY 721 GAGAAGATCTAGGCTGTTGTACCAAGCACTACACAGGGAATTCACCTGATCGAG 780  
Db 721 GAGAAGATCTAGGCTGTTGTACCAAGCACTACACAGGGAATTCACCTGATCGAG 780  
OY 781 GTAAGATTCACCTGGAAGGCGAGATGGCTACTATCTGATGATGATGATGATGATG 840  
Db 781 GTAAGATTCACCTGGAAGGCGAGATGGCTACTATCTGATGATGATGATGATGATG 840  
OY 841 CTACTCATGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900  
Db 841 CTACTCATGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900  
OY 901 CGTGTGGGCTGGGCTGCAACACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960  
Db 901 CGTGTGGGCTGGGCTGCAACACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960  
OY 961 GCGCTTTGGCTGAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020  
Db 961 GCGCTTTGGCTGAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020  
OY 1021 CTTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080  
Db 1021 CTTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080  
OY 1081 AAGAATTCATGAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1140  
Db 1081 AAGAATTCATGAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1140  
OY 1141 CAAGAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200  
Db 1141 CAAGAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200  
OY 1201 GAGGATTCATGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260  
Db 1201 GAGGATTCATGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260  
OY 1261 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1320  
Db 1261 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1320

RESULT 2  
US-10-075-846-1  
; Sequence 1, Application US/10075846  
; Publication No. US20030032608A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING A NOVEL GLYCINE RECEPTOR ALPHA SUBUNIT  
; FILE REFERENCE: D0079 NP  
; CURRENT APPLICATION NUMBER: US/10/075,846  
; PRIOR FILING DATE: 2002-02-13  
; PRIOR APPLICATION NUMBER: US 60/269,535  
; NUMBER OF SEQ ID NOS: 81  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1  
; LENGTH: 2565  
; TYPE: DNA  
; ORGANISM: homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1251)  
US-10-075-846-1

Query Match 93.0%; Score 1524.8; DB 9; Length 2565;  
Best Local Similarity 97.3%; Pred. No. 0;  
Matches 1578; Conservative 0; Mismatches 2; Indels 42; Gaps 1;  
OY 1321 CGGGCTGTCTCCCTTTCACCTTTCATCTTCATATCTTCTACTGGGTGCTATAAA 1380  
Db 1321 CGGGCTGTCTCCCTTTCACCTTTCATCTTCATATCTTCTACTGGGTGCTATAAA 1380  
OY 1381 GTGCTATGTGTCAGAGATATCCACAGCGCTGTGTGAATAGGGTGAAGTATAGACTCT 1440  
Db 1381 GTGCTATGTGTCAGAGATATCCACAGCGCTGTGTGAATAGGGTGAAGTATAGACTCT 1440  
OY 1441 GCTGCTGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500  
Db 1441 GCTGCTGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500  
OY 1501 GTTGGACAGTTCCTTCTGATCTCCACTCAGAACTTCACTACAGTCCCAAGCTAT 1560  
Db 1501 GTTGGACAGTTCCTTCTGATCTCCACTCAGAACTTCACTACAGTCCCAAGCTAT 1560  
OY 1561 GTGGGCTATATGTCATGTCGTCATGTCGTCATGTCGTCATGTCGTCATGTCGTC 1620  
Db 1561 GTGGGCTATATGTCATGTCGTCATGTCGTCATGTCGTCATGTCGTCATGTCGTC 1620  
OY 1621 TAAAAAATGAAAAAATGAAAAA 1640  
Db 1621 TAAAAAATGAAAAAATGAAAAA 1640



```
OY 301 CAGTGAATGACCCAGCGCTGCTACCGAGAAATCTGTGATGACTCTGTGGACCTGCAT 360
DB 301 CAGTGAATGACCCAGCGCTGCTACCGAGAAATCTGTGATGACTCTGTGGACCTGCAT 360
OY 361 CCCTCCATGCTGGAGCTCTATCTGGAAGCCAGACCTCTTCTTGGCTAATGAGAAAGGGCC 420
DB 361 CCCTCCATGCTGGAGCTCTATCTGGAAGCCAGACCTCTTCTTGGCTAATGAGAAAGGGCC 420
OY 421 AACCTCCATGAGAGTGACCGAGCAACAAGTACTGCGGCATCTTCAAAATGGGAATGTG 480
DB 421 AACCTCCATGAGAGTGACCGAGCAACAAGTACTGCGGCATCTTCAAAATGGGAATGTG 480
OY 481 CTGTACAGCATCAGGCTGACCCCTCATTTTGTCTGCTGATGAGACCTCAAGAACTTCCCC 540
DB 481 CTGTACAGCATCAGGCTGACCCCTCATTTTGTCTGCTGATGAGACCTCAAGAACTTCCCC 540
OY 541 ATGACATCCAGAGCTGACAGATGACAGCTTGAAGACCTCATCTGACAGCCCTCTG 600
DB 541 ATGACATCCAGAGCTGACAGATGACAGCTTGAAGACCTCATCTGACAGCCCTCTG 600
OY 601 CCATCTCTGTCACTTTCAGTTGGCTACACCATGAAAGACCTGCTGTGAGTGGCTGGA 660
DB 578 -----TTGGCTACACCATGAAAGACCTGCTGTGAGTGGCTGGA 618
OY 661 GATGCTCCTGCTGTCACAGTGAGGAGGCTGACTGCGCCAGTTTATCTTGGCGGAT 720
DB 619 GATGCTCCTGCTGTCACAGTGAGGAGGCTGACTGCGCCAGTTTATCTTGGCGGAT 678
OY 721 GAGAAGATCTAGGCTGTTGTACCAAGCATCAACACAGGGAATTCACCTGCTGACG 780
DB 679 GAGAAGATCTAGGCTGTTGTACCAAGCATCAACACAGGGAATTCACCTGCTGACG 738
OY 781 GTAAAGTTTCACTGGAAGCGGACATGGGCTACTATCTGATTCAGATGTACATCCCGAC 840
DB 739 GTAAAGTTTCACTGGAAGCGGACATGGGCTACTATCTGATTCAGATGTACATCCCGAC 798
OY 841 CTACTCATGCTCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
DB 799 CTACTCATGCTCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 858
OY 901 CGTGTGGGCTGGGCGATCACCAGCGTCTCATCATGACACCCAGAGCTTGGCTCCGG 960
DB 859 CGTGTGGGCTGGGCGATCACCAGCGTCTCATCATGACACCCAGAGCTTGGCTCCGG 918
OY 961 GCCTCTTGGCTTAAAGGTGCTTCACTGGAAGCAATCGACATCTGATGGCTGTGCTG 1020
DB 919 GCCTCTTGGCTTAAAGGTGCTTCACTGGAAGCAATCGACATCTGATGGCTGTGCTG 978
OY 1021 CTCTTTGTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
DB 979 CTCTTTGTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1038
OY 1081 AAAGCATTCATACGACTTCGAAAGGAGGAGGCGCCAAACGCTTGGAGAGATATCATC 1140
DB 1039 AAAGCATTCATACGACTTCGAAAGGAGGAGGCGCCAAACGCTTGGAGAGATATCATC 1098
OY 1141 CAAGAAATCGTCTTATTTCCGTGCTATGCTTGGGCGCACTGCTGCAAGCAAGAT 1200
DB 1099 CAAGAAATCGTCTTATTTCCGTGCTATGCTTGGGCGCACTGCTGCAAGCAAGAT 1158
OY 1201 GGAGTCCCAATGGAAGTTCTGCGATTTATAGTCCCAACCTCAGGCCCTTCTTAAG 1260
DB 1159 GGAGTCCCAATGGAAGTTCTGCGATTTATAGTCCCAACCTCAGGCCCTTCTTAAG 1218
OY 1261 GAAGAGAAACACGCGGAAAACCTTACGTGACCTGAGCCAAAGAAATTTGACACATCTCC 1320
DB 1219 GAAGAGAAACACGCGGAAAACCTTACGTGACCTGAGCCAAAGAAATTTGACACATCTCC 1278
OY 1321 GGGGCTGCTTCCCTTCACTTCTCATCTTCAATATCTTCACTGGGCTGTCTATAA 1380
DB 1279 GGGGCTGCTTCCCTTCACTTCTCATCTTCAATATCTTCACTGGGCTGTCTATAA 1338
```

```
OY 1381 GTGCTATGTCAGAGATATATCACCAGCGCTGTGTAATAGGGTGAGACTATAGACTCT 1440
DB 1339 GTGCTATGTCAGAGATATATCACCAGCGCTGTGTAATAGGGTGAGACTATAGACTCT 1398
OY 1441 GCTGCTGCGCTCTGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
DB 1399 GCTGCTGCGCTCTGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1458
OY 1501 GTTGGACAGTTCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560
DB 1459 GTTGGACAGTTCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1518
OY 1561 GTGGCTATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
DB 1519 GTGGCTATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1578
OY 1621 TA 1622
DB 1579 TA 1580

RESULT 3
US-10-075-846-9
; Sequence 9, Application US/10075846
; Publication No. US20030032608A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING A NOVEL GLYCINE RECEPTOR ALPHA SUBUNIT
; FILE REFERENCE: D0079 NP
; CURRENT APPLICATION NUMBER: US/10/075,846
; PRIOR FILING DATE: 2002-02-13
; PRIOR APPLICATION NUMBER: US 60/269,535
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 993
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-075-846-9

Query Match 49.3%; Score 807.8; DB 9; Length 993;
Best Local Similarity 90.0%; Pred. No. 1.9e-253;
Matches 931; Conservative 0; Mismatches 2; Indels 102; Gaps 2;

OY 1 ATGACACTCTTGTCTCTGCAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 60
DB 1 ATGACACTCTTGTCTCTGCAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 60
OY 61 GTCCCTCTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
DB 61 GTCCCTCTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
OY 121 CCCATGCTCCCTCTGATTTCTTATAGCAAACTTATGAGGCGCAACATCTGATATGATGCC 180
DB 121 CCCATGCTCCCTCTGATTTCTTATAGCAAACTTATGAGGCGCAACATCTGATATGATGCC 180
OY 181 AGGATTTGGCCCAATTTTAAAGGCCACCCGTTGAAGCTGACCTGCAACATCTTCAAC 240
DB 181 AGGATTTGGCCCAATTTTAAAGGCCACCCGTTGAAGCTGACCTGCAACATCTTCAAC 240
OY 241 AGTTTCAGCTCCGCTACCAAGCAACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 269
DB 241 AGTTTCAGCTCCGCTACCAAGCAACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
OY 270 -----GGACTACCGGCTGAATGCTCTTCTGCGGCA 300
DB 301 TCTGAAGGCCCATATCTGCAACCTTCCAGAGACTACCGGCTGAATGCTCTTCTGCGGCA 360
OY 301 CAGTGAATGACCCAGCGCTGCTACCGAGAAATATCTGATGACTCTGTGACCTGCAT 360
DB 361 CAGTGAATGACCCAGCGCTGCTACCGAGAAATATCTGATGACTCTGTGACCTGCAT 420
```

Db	193	AATTTCCAGGAGGAAGAATATCATCCAAAGAGTGTTTCTATTTCCGGGCTATAGGCTGG	252
QY	1178	GCCACTCCTCGCAGAGCAGAGATGAGAGTCCAAATGGAAGTTCTGGCATTTATATGCC	1237
Db	253	GCCACTCCTCGCAGAGCAGAGATGAGAGTCCAAATGGAAGTTCTGGCATTTATATGCC	312
QY	1238	AACCTCCAGGCCCTCTCTAAGGGAAAGAGAAACACGGGAAACCTACGTGACTGAG	1297
Db	313	AACCTCCAGGCCCTCTCTAAGGGAAAGAGAAACACGGGAAACCTACGTGACTGAG	372
QY	1298	CCAGAGAAATTTGACACCATCTCCCGGGCTGTCTTCCCTTACATTCCATCTTCAATA	1357
Db	373	CCAGAGAAATTTGACACCATCTCCCGGGCTGTCTTCCCTTACATTCCATCTTCAATA	432
QY	1358	TCCTTCTACTGGGTTGTCTATPAAGTCTATGTCAGAAATATCCACAGGCTGTGAA	1417
Db	433	TCCTTCTACTGGGTTGTCTATPAAGTCTATGTCAGAAATATCCACAGGCTGTGAA	492
QY	1418	TAGGGTGGAGACTATTAAGTCTGCTGTGGCCCTCGATCCGTCCTGGGTGGGCTTCT	1477
Db	493	TAGGGTGGAGACTATTAAGTCTGCTGTGGCCCTCGATCCGTCCTGGGTGGGCTTCT	552
QY	1478	CCCTCAGTTAGACTCATTAAGGGGTTTGACAGATTCTTCCTGATCTCCACTCAGAACT	1537
Db	553	CCCTCAGTTAGACTCATTAAGGGGTTTGACAGATTCTTCCTGATCTCCACTCAGAACT	612
QY	1538	TCAACTACCACTCCCAAGCATATGTGGGCTCATATTTCATATGTCGCCAATGGTGGCTTAC	1597
Db	613	TCAACTACCACTCCCAAGCATATGTGGGCTCATATTTCATATGTCGCCAATGGTGGCTTAC	672
QY	1598	TTATATAA 1604	
Db	673	TTATATAA 679	

RESULT 5

US-09-918-995-15843

Sequence 15843, Application US/09918995

Publication No. US20030073623A1

GENERAL INFORMATION:

APPLICANT: Hyseq, Inc.

TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED

TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES

FILE REFERENCE: 20411-756

CURRENT APPLICATION NUMBER: US/09/918, 995

CURRENT FILING DATE: 2001-07-30

PRIOR APPLICATION NUMBER: US/09/235, 076

PRIOR FILING DATE: 1999-01-20

NUMBER OF SEQ ID NOS: 38054

SOFTWARE: FASTSEQ for Windows Version 3.0

SEQ ID NO 15843

LENGTH: 492

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc\_feature

LOCATION: (1)...(492)

OTHER INFORMATION: n = A,T,C or G

US-09-918-995-15843

Query Match	Best Local Similarity	13.4%: Score 219; DB 9; Length 492;
Matches 219; Conservative	100.0%; Pred. No. 6,8e-61;	
	0; Mismatches 0; Indels 0; Gaps 0;	
QY	1404	CCAGGCTCTGTAATAGAGGTGGAGACTTAAAGTCTCTGCTGCGCTCTGCTTCTCC
Db	57	CCAGGCTCTGTAATAGAGGTGGAGACTTAAAGTCTCTGCTGCGCTCTGCTTCTCC
QY	1464	TGGGTGGGCTTCTCCCTCAGTTAGAGTCCATTAAGGGTTTGGAGAGTTCTTCTGATC
Db	117	TGGGTGGGCTTCTCCCTCAGTTAGAGTCCATTAAGGGTTTGGAGAGTTCTTCTGATC
QY	1524	TCCCACTCAGAACTTCAACTACCACTCCCAAGCATATGTGGGCTCATATTGCATATGTC
Db	1583	TCCCACTCAGAACTTCAACTACCACTCCCAAGCATATGTGGGCTCATATTGCATATGTC

|||||  
Db 177 TCCACCTCAGAACTCACTACACAGTCCCAAGCTATGTGGCCTATATGACATGTGCC 236  
Oy 1584 AATGGTGGCTGTACTTATAAGATGGCTTATCTACCTA 1622  
Db 237 AATGTGTGTACTTATTAAGATGGCTTATCTACCTA 275

RESULT 6  
US-10-211-673-11

Sequence 11, Application US/10211673  
Publication No. US20030013158A1  
GENERAL INFORMATION:  
APPLICANT: Le Bourdellies, Beatrice  
TITLE OF INVENTION: HUMAN ALPHA 4 RECEPTOR SUBUNIT  
OF THE GABA-A RECEPTOR  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merck & Co., Inc.  
STREET: P.O. Box 2000, 126 E. Lincoln Ave.  
City: Rahway  
STATE: NJ  
COUNTRY: USA  
ZIP: 07065-0900  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
OPERATING SYSTEM: Windows  
SOFTWARE: FASTSEQ for Windows Version 2.0b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/211.673  
FILING DATE: 02-Aug-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB95/02323  
FILING DATE: 29-SEP-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Xu, Yang  
REGISTRATION NUMBER: 45,243  
REFERENCE/DOCKET NUMBER: T1292  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 732-594-1307  
TELEFAX: 732-594-4720  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1555 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: Coding Sequence  
LOCATION: 47...1402  
OTHER INFORMATION:  
SEQUENCE DESCRIPTION: SEQ ID NO: 11:  
US-10-211-673-11  
Query Match 11.8%; Score 193; DB 9; Length 1555;  
Best Local Similarity 53.3%; Pred. No. 4,5e-52;  
Matches 486; Conservative 0; Mismatches 380; Indels 45; Gaps 2;

Db 287 GCCAATGAGTACACCATGACGGTGTCTTCGACACGAGACTGGCGGAGACAGGCTC 346  
Oy 322 TCCACGAGAAATATCTGTATGACTCTCTGACCTGCATCCCTCCATCGTACACTATC 381  
Db 347 TCCATAC---AACCAACACCAAGAGACCTGGGCTTGACAGACCGCTTGTGGCAACCTG 403  
Oy 382 TGGAAAGCCAGACCTCTTCTTGTGTAATGAGAAAGAGGCGCAACTTCATGAGGTACACG 441  
Db 404 TGGCTGCCGACACCTTCTATCTGTGAACGCCAAGTGGGCTGTTCACAGACGTACAGGTG 463  
Oy 442 GACAAACAAGTTACTGCGCATCTTCAGAAATGGAAATGCTGTACACATCAGGCTGACG 501  
Db 464 GAGAAACAAGTCAATCCGCTGACAGCCGAGGGGTGATCTGTACACATCCGAATACG 523  
Oy 502 CTCAATTTGTCTTCCCTGATGAGACCTCAAGACTTCCCATGGACATCCAGAGCTGACG 561  
Db 524 TCCACTGTGCTCCGACATGAGACCTGCGCAAAATTCCTCCATGACGAGAGAGTGTGATG 583  
Oy 562 ATGAGCTTGAGACCTCATCATCTGACAGCCCTCTGCCATCTGTCACTTTCAGTT 621  
Db 584 CTGGACCTGGAGACT-----AC 601  
Oy 622 GCGTACACATGAAGAAGCTGTGTGTTGAGTGCCTGAGAAATGCTCTGCTGTCCAAATG 681  
Db 602 GGTACTCATCGAGACATCGTCTACTAGTGTGGAAGACGAGACATCCACGAGG 661  
Oy 682 GCTGAGGGCTGACTCTGCCCATTTATCTTGGGGATGAGAAGATCAAGCTGTGT 741  
Db 662 CTGACAAAGCTGACGCTGGCGCAGATTCACCATCACGATACCGCTTACACGAGAGCTG 721  
Oy 742 ACCAAGCATTACAAACACAGGAAATTCACCTGCATCGAGGTAAAGTTTCACTGGAACGG 801  
Db 722 ATGAACCTTAAAGTCCGCTGGCGCAATTCACAGGCTGACGCTGCACCTTCACCTGGGAG 781  
Oy 802 CAGATGGCTACTATCTGATTCAGATGTACATCCCAAGCCTACTCATCTCATCTGTCC 861  
Db 782 AACCGCGCGTGTACATCATCAATCTACATCGCTCCGCTGCTGCTGCCATCTCC 841  
Oy 862 TGGTCTCTCTTGTGATCAACATGATGATGCGCCCTCCGCTGGGCGCTGGGATACG 921  
Db 842 TGGTCTCTCTTGTGATGACGACGAGCGCGGCTGCCGCGAGGCTGTCTTGAAGATACG 901  
Oy 922 ACCGTCTCAACATGACACACACAGAGCTGTGGCTCCCGGCGCTTGTGCTTAAGGTGTC 981  
Db 902 ACGGTCTGAGATGACACAGAGCTATGTGATGCTCCGCTCTCCGACAGGCAATCA 961  
Oy 982 TACGTGAAGCAATGACATCTGTGATGCTGTGTCTCTTGTGTTGCTGCTGCTG 1041  
Db 962 GCCATCAAGGACGACGATCTACTTCTGTGATCTGTATGTCTGTGTTGGCGGCTG 1021  
Oy 1042 CTGAGATATGC 1052  
Db 1022 GTGAGATACG 1032

RESULT 7  
US-10-239-420-10

Sequence 10, Application US/10239420  
Publication No. US20030096984A1  
GENERAL INFORMATION:  
APPLICANT: Cully, Doris F.  
APPLICANT: Zheng, Yingcong  
TITLE OF INVENTION: DNA MOLECULES ENCODING LIGAND GATED ION  
CHANNELS FROM DERMACEPTOR VARIANTS  
FILE REFERENCE: 20629P  
CURRENT APPLICATION NUMBER: US/10/239.420  
CURRENT FILING DATE: 2002-09-23  
PRIOR APPLICATION NUMBER: PCT/US01/09956  
PRIOR FILING DATE: 2001-03-28  
PRIOR APPLICATION NUMBER: 60/193,935  
PRIOR FILING DATE: 2000-03-31  
NUMBER OF SEQ ID NOS: 13  
SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 10  
LENGTH: 1197  
TYPE: DNA  
ORGANISM: Rhiplocephalus sanguineus  
US-10-239-420-10

Query Match 11.4%; Score 187.6; DB 9; Length 1197;  
Best Local Similarity 55.0%; Pred. No. 2.3e-50;  
Matches 466; Conservative 0; Mismatches 334; Indels 48; Gaps 3;

226 AACACCTCATCAACAGTTTCCGCTCCGCTACCAAGACACAAAGACTACGGGTAT 285  
Db AACATCTTTGATAGATGAGCAATGATGATGATGATGATGATGATGATGATGAT 156  
Oy AACATCTTTGATAGATGAGCAATGATGATGATGATGATGATGATGATGATGAT 345  
Db ATGACGTTTCAGAGCACTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 216  
Oy TCTCTGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 405  
Db GTTCGCTACCTGACGCTGACGAGACCGAGACAGCTTGGAGCGGAGCTTTTCTCC 276  
Oy AATGAGAAAGGGGACCACTTCCATGAGGTGACCAAGCAACAAGTTACTGGGATCTTC 465  
Db AACGAGAAAGAGGACACTTCCAAACATCATCATGCCCCAGTCTCTTACGATACAT 336  
Oy AAGATGGGAATGCTGCTAGACATCAGGCTGACGCTCATCTTCTCTGCTGATGAGAC 525  
Db CCCAAGGGGAGAGCTCTCTTACGATGAGAAATATCTTGGTCTTCAATGTCGATGAG 396  
Oy CTCGAAACCTTCCCATGAGATCCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 585  
Db CTGAAATTTTATCTTTGATTAACAAATCTGCTTATGCTATGCTAGCTGAGCT----- 448  
Oy CTCGACAGCCCTGCTCCATCTCTGCTATGCTATGCTATGCTATGCTATGCTATGCT 645  
Db -----ATGGGTATACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 474  
Oy TTTGATGGCTGAGAGATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 705  
Db TTTTATG---GAAAGAGGGGAGCTCTGATGAGTCAAAAATCTTCACTTGGCAGCT 531  
Oy TTTATCTTGGGAGATGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 765  
Db TTTCA---CGCTGGAAGGTTTCAAAACGACTACTGACAGAGTGGAGCAACACTGGCAG 588  
Oy TTTACCTGATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 825  
Db TTTACCTGATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 648  
Oy ATGTACATCCCGAGCTACTACTGATGATGATGATGATGATGATGATGATGATGATG 885  
Db ATGTACATCCCGAGCTACTACTGATGATGATGATGATGATGATGATGATGATGATG 708  
Oy GATGCTGCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 945  
Db ACCGTGATCCCGGCGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 768  
Oy AGCTTGGGCTCCCGGCTCTTGGCTTAAGTGTCTGATGATGATGATGATGATGATGAT 1005  
Db ATATGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 828  
Oy ATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1065  
Db ACCGGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 888  
Oy 1066 GTTTCTCG 1073  
Db 889 GCCTCGCG 896

RESULT 8  
US-09-969-844-13

Sequence 13, Application US/09969844  
Publication No. US2002019276A1  
GENERAL INFORMATION:  
APPLICANT: Xiao-Zhou Michelle Wang  
APPLICANT: Xavier Georges Sarda  
APPLICANT: Michael David Tomalski  
APPLICANT: Vincent Paul Mary Wingate  
TITLE OF INVENTION: Heliothis glutamate Receptor  
FILE REFERENCE: A32815-1 07267.0178  
CURRENT APPLICATION NUMBER: US/09/969, 844  
CURRENT FILING DATE: 2001-10-03  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 13  
LENGTH: 4621  
TYPE: DNA  
ORGANISM: Heliothis virescens  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (764)...(764)  
OTHER INFORMATION: n is a, t, g, or c  
US-09-969-844-13

Query Match 11.4%; Score 187; DB 9; Length 4621;  
Best Local Similarity 54.5%; Pred. No. 8.2e-50;  
Matches 472; Conservative 0; Mismatches 346; Indels 48; Gaps 3;

Oy CCCGTAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 267  
Db CCAGCGGAGTATGAGGCTCAATATATTTGTCGAGATATATCAAGATGATGATGATGAT 389  
Oy ATGACTACCGGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 327  
Db ATGGAATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 449  
Oy CGAGATATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 387  
Db AATATCTTGGAGAGTGGCTCAATATCTGATGATGATGATGATGATGATGATGATGAT 509  
Oy CCAGACCTCTTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 447  
Db CCGATCTATCTTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 569  
Oy AAGTACGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 507  
Db GTGTACATCCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 629  
Oy TTTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 567  
Db CTCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 689  
Oy CTTGAGACCTCATCACTACTGATGATGATGATGATGATGATGATGATGATGATGATG 627  
Db ATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 707  
Oy ACCATGAAAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 687  
Db ACCAGAGAGAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 764  
Oy GGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 747  
Db AACTTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 821  
Oy CACTACACACAGGAGAAATCACTGATGATGATGATGATGATGATGATGATGATGATG 807  
Db AAGACTAATATCCGGAATATGATGATGATGATGATGATGATGATGATGATGATGATG 881  
Oy GGCTACTATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 867  
Db ACTTACTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 941  
Oy TCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 927



```
Db 477 CAACGAGAGGAGGATCTTCCACAAATCATCATCCGGAACCTGTATACCCGATCTT 536
Oy 465 CAAGATGGGAATGCTGTACACATCAGGCTGACCCCTCTTTTGTCCCTGGATGGA 524
Db 537 CCCCAGGAGGACGCTGTACAGATCCGATCTCCCTGACGCTCTGCTGCCCAATGAA 596
Oy 525 CCTCAAGAACTTCCCATGACATCCAGACGCTGACAGATGACGCTTGAAGCTCATCAT 584
Db 597 CCTCAAGTGTATACCCCGATTAAGACAGACTGCTGCTGATAGATGCTTACTT----- 649
Oy 585 ACTTCGACGCCCTTCCCATCTGTCTACTTTCAGTTGGCTACACCATGAAGACCTGCT 644
Db 650 -----ATGTTGGACCAACAGACGACTTATG 674
Oy 645 GTTTAGTGGCTGGAAGATGCTCTGCTGTCCAGATGCTGAGGAGGCTGACTTGGCCCA 704
Db 675 GTTCTATGGAAGGAGGCGAC---CGGTGAGGTGTGTAAGAACTTACACCTGCTCG 731
Oy 705 GTTATCTTGGGGATGAGAAGATCTAGGCTGTTGTACCAAGCACTACACAGAGGAA 764
Db 732 GT---TACGCTGGAGAGTCTCTACACTAGCTACGCAACAGTAAAGTAAATACCGGTGA 788
Oy 765 ATTCACCGCATCGAGGTAAGTTTCACTGGAACGGGAGATGGGCTACTATCTGATTC 824
Db 789 ATACAGTTGCTGAAGGTAGACCTGCTCTTCAAGCGAGTTCACTTACTACCTGATCCA 848
Oy 825 GATGATACATCCCGACCTACTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 884
Db 849 GATGATACATTCGAGCTGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 908
Oy 885 GATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 944
Db 909 GGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 968
Oy 945 GAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1004
Db 969 GTCTGACGAGCATCAACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1028
Oy 1005 GATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1064
Db 1029 GACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1088
Oy 1065 TGTCTCTCG 1073
Db 1089 TCGCTCTCG 1097

RESULT 11
US-10-239-420-3
; Sequence 3, Application US/10239420
; Publication No. US2003009684A1
; GENERAL INFORMATION:
; APPLICANT: Cully, Doris F.
; APPLICANT: Zheng, Yingcong
; TITLE OF INVENTION: DNA MOLECULES ENCODING LIGAND GATED ION
; TITLE OF INVENTION: CHANNELS FROM DERMACEPTOR VARIANTS
; FILE REFERENCE: 20629P
; CURRENT APPLICATION NUMBER: US/10/239,420
; CURRENT FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: PCT/US01/09956
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: 60/193,935
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 3442
; TYPE: DNA
; ORGANISM: Dermaceptor variabilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (32)...(1225)
US-10-239-420-3
```

```
Query Match 10.5%; Score 171.8; DB 9; Length 3442;
Best Local Similarity 52.9%; Pred. No. 6.5e-45;
Matches 454; Conservative 0; Mismatches 357; Indels 48; Gaps 2;

Oy 202 GGGCCACCCGTGAGACGACCTGCAACATCTTCAATCAACAGTTTCACCTCCGTACCAAG 261
Db 197 GGAAGGCCAACMAAGAGGCTTGGAAATCTACATACGACCTTGGGTCCATTAATCCA 256
Oy 262 ACCAATATGACCTACCGGATGATGTCTTCTTGGGGACAGAGTGAAGTGAAGCCAGCGCTG 321
Db 257 GCCAATATGACCTATGAGATGTGATCTTATTTTGGGGACAGCTTGGCAAGATGATCGCTTG 316
Oy 322 TCTACCGAGAATATCTGTATGATGACTCTGTGACCTCGATCCCTCATCTGAGCTATC 381
Db 317 ACAGGCCCAACGATATCCAGGCCCCGTGACCTCAATGATCCAAAGCTGTGACAGCTATA 376
Oy 382 TGAAGCCAGACCTCTTCTTGTATGAGAAAGGGCCAACTTCCATGAGTGAACCAAG 441
Db 377 TGAAGCCAGAGATCTTCTTGTGCAAAAGCAACAGAGGTTCCATATGATGACAGTA 436
Oy 442 GACAAACAGTACTGCGCATCTTCAGAAATGGAATGCTGTACAGATCAGGCTGACC 501
Db 437 CCTATATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 496
Oy 502 CTGATTTTGTCTGCTGCTGATGAGCTCAAGACTTCCCATGACATCCAGAGCTGACG 561
Db 497 CTGAGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 556
Oy 562 ATGACAGCTGAGAGCTCATCTGATGCTGAGAGCCCTGCTGCTGCTGCTGCTGCTGCTG 621
Db 557 ATGACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 580
Oy 622 GGTACACCATGAAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 681
Db 581 -----ACACCGAAGAACTGATGAGAGTGTCCGATPACCAATCCGATTAATCTA 631
Oy 682 GCTGAGGGGCTGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 741
Db 632 TTTGAGAGGCTGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 688
Oy 742 ACCAAGCTACACACAGGGAATTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 801
Db 689 ATGAGAAATTTTACATGCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 748
Oy 802 CAGATGGCTACTATCTGATTCAGATGATATCCCGACCTACTCATCTGCTGCTGCTGCTG 861
Db 749 TCACTGGGCTACACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 808
Oy 862 TCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 921
Db 809 TGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 868
Oy 922 ACCGTGCTACCATACCAAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 981
Db 869 ACCGTGCTACATTTTCTTCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 928
Oy 982 TACGTGAAGCAATGACATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1041
Db 929 TACGTGAAGCAATGATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 988
Oy 1042 CTGAGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1060
Db 989 CTGAGATTCACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1007

RESULT 12
US-09-815-925-10
; Sequence 10, Application US/09815925
; Patent No. US20020127199A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
```

```

: APPLICANT: Goodrich, Ryle
: APPLICANT: Asundi, Vinod
: APPLICANT: Yang, Yonghong
: APPLICANT: Zhang, Jie
: APPLICANT: Wehrman, Tom
: APPLICANT: Drmanac, Radote T.
: TITLE OF INVENTION: No. US20020127199A1el Nucleic Acids and
: TITLE OF INVENTION: Polypeptides
: FILE REFERENCE: 787CIP2H
: CURRENT APPLICATION NUMBER: US/09/815,925
: PRIOR FILING DATE: 2001-03-22
: PRIOR APPLICATION NUMBER: 09/560,875
: PRIOR FILING DATE: 2000-04-27
: PRIOR APPLICATION NUMBER: 09/496,914
: PRIOR FILING DATE: 2000-02-03
: NUMBER OF SEQ ID NOS: 11
: SOFTWARE: PL_FL_genes Version 2.0
: SEQ ID NO 10
: LENGTH: 1467
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (46)..(1464)
: US-09-815-925-10

```

Query Match 10.3%; Score 168.6; DB 10; Length 1467;  
Best Local Similarity 52.6%; Pred. No. 4.3e-44;

Matches 474; Conservative 0; Mismatches 379; Indels 48; Gaps 3;

```

OY 169 GGATATGATGCCAGATTGCGCCCAATTTAAAGCCCAACCGGTGACGTGACCTGCAAC 228
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 184 GGTAGACGATTCGCGCTAAGACCGGACTTCGGGGGTCGCCCGCTGCTGGGGGATGAAC 243
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 229 ACTTCATCAAGCTTCCAGTCCGACCAAGACACCAATGAGTACCGGGGATGTC 288
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 244 ATGACATGCGCCGACATGACATGTTCCGAGTCAACATGATTAACCTTAACCATG 303
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 289 TTTCTGCGGCAAGTGAATGACCAACCGCTGCTCCAGCAATATCTGATGACTCT 348
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 304 TATTTCAACAATATTTGAGAGATTAAGAGTCCCTATTTCTGGATCCCTCCAC--- 360
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 349 CTGGACCTGATCCCTCCATGCTGAGCTATCTGAGACCCAGACCTCTTTTGTCTAAT 408
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 361 CTCACCTTGACATGAGTGGCTGACACAGCTATGGTCCCGACACATATTTCTTAAT 420
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 409 GAGAAAGGGGCCAATTCATGATGATGACACGACGACCAAGCAAGTACTGGGCATCTCAAG 468
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 421 GACAAAAGTCATTTGTGATGATGAGTGAAGTAAACCGCATGATCCGCTTTCACCT 480
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 469 AATGGATATGCTGTACAGCATCAGGCTGACCCCTCATTTTGTCTGCTGATGACCTC 528
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 481 GATGGACAGTGTGTATGAGGCTCAGATTCACACACACACACACACATGATGAGCTC 540
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 529 AAGACTTCCCATGACATCCAGCTGACAGCTGACCTGAGAGCTCATCTCATCTC 588
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 541 AGGAGTACCCCTCCGACGACGAGAACTGACCTGCTGAAATGAAAGCT----- 589
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 589 TGCAGCCTCTGACATCTGTGACATTTGATGATGACCAATGAAAGCACTGCTGTT 648
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 590 -----ATGGCTACACACGAGTGAATGATGATGTTGTT 618
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 649 GAGTGGCTGGAAGATCTCTGCTGTCAAGTGGCTGAGGGGCTGACTGCTGCCACAGTT 708
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 619 TACTGCGAGGCGGGGCAAGGCTTTACCGGAGTGAAGATGAGTCCCGGAGT-- 676
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 709 ATTTCCGGATGAGAAGATCTAGCTGTTTACCAAGACTACAAACAGGAAATTC 768
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 677 -TCTTCATGATGAGACCGCTGCTGAGAAATGTTCTTCCACAGGTGCTAT 735
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 769 ACCTGACGAGTAAAGTTTCACTGGAACGACAGATGGGCTACTATGATTCAGATG 828
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 736 CCTGCACTGTCACTGACCTTTCGTTGAAGAGAACTGATGATCTTCACTTTCAGACT 795
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

OY 829 TACATCCCAAGCCTACTCATCTGCTGCTGCTGCTCTTGTGATCAACATGAT 888
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 796 TATATCCCTTAATGATGATTAAGATCTGTGTGGGTCTCTTGTGATCAATATGAT 855
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 889 GCTGCCCTTCCGCTGTGGGCTGAGCATCACCGCTCTACCATGACCAACGAGAC 948
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 856 GATCTGCTGCTAGAGTCTCCCGGATCACAACTGTCTGATCAACCAACATCAAC 915
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 949 TCTGCTCCCGGCTCTTGGCTAAGTGTCTGATGAGCAATGACATCTGGATG 1008
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 916 ACCCACTTTCGGAGACTTGGCCCAAAATCCCTATGTCAAAAGCCATGATGACCTT 975
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1009 GCTGTGCTCTGCTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1068
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 976 ATGGGCTGCTGCTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1035
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1069 T 1069
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1036 T 1036
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 13  
US-10-239-420-1  
Sequence 1, Application US/10239420  
Publication No. US20030096984A1  
GENERAL INFORMATION:

```

: APPLICANT: Cully, Doris F.
: APPLICANT: Zhang, Yingsong
: TITLE OF INVENTION: DNA MOLECULES ENCODING LIGAND GATED ION
: TITLE OF INVENTION: CHANNELS FROM DERMACENTOR VARIABILIS
: FILE REFERENCE: 20629P
: CURRENT APPLICATION NUMBER: US/10/239,420
: PRIOR FILING DATE: 2002-09-23
: PRIOR APPLICATION NUMBER: PCT/US01/09956
: PRIOR FILING DATE: 2001-03-28
: PRIOR APPLICATION NUMBER: 60/193,935
: PRIOR FILING DATE: 2000-03-31
: NUMBER OF SEQ ID NOS: 13
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1
: LENGTH: 3598
: TYPE: DNA
: ORGANISM: Dermacentor variabilis
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (170)..(1363)
: US-10-239-420-1

```

Query Match 10.2%; Score 167; DB 9; Length 3598;  
Best Local Similarity 52.5%; Pred. No. 2.5e-43;

Matches 451; Conservative 0; Mismatches 360; Indels 48; Gaps 2;

```

OY 202 GGGCCACCGGTGAAGCTGACATCTGCAATCTTCATCAACAGTTTCACTCCGACCAAG 261
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 335 GGAAAGCCCAAAAGATGCTGCGAAATCTACAGGCTCGCGGTGCAATAATCCA 394
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 262 ACCCAATGAGTACAGCGGATGATGCTTCTGCGGCAACAGTGGAAATGACCCACCG 321
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 395 GCCCAATGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 454
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 322 TCTTACAGGAATATCTGATGATGATGATGATGATGATGATGATGATGATGATG 381
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 455 ACAGAGCCCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 514
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 382 TGAAGCCAGACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 441
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 515 TGAAGCCAGGAATATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 574
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 442 GACAAACAGTACTGCTGATCTTCAAGAAATGGAATGTGCTGATGATGATGATGATG 501
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 575 CCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 634
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```



Oy	502	CTCATTTTGTGCTGCTGATGAGACTCAGAAACTTCCCATGAGATCCAGACGTGCACG	561
Db	635	CTAAGGTTTGATGATGATGATGATTTAATATCGCTTTCTATGAGACTCCCAAGTTTGACG	694
Oy	562	ATGCAAGCTTGAAGAGCTCATCTCATACTCTGCAGACCCTCTGCATCTCTGTCACTTTCAGTT	621
Db	695	ATGCAGACTCGGCTCATCTATCTCGAAA-----	718
Oy	622	GGCTACACCATGAAAGACCTCGTGTGTTGAGTGGCTGGAAGATGCTCTGTGTCAGTG	681
Db	719	-----ACAAACCGAAGAACTGCTATGAGTGGTCTGATACCAATCCGATATACTA	769
Oy	682	GCTGAGGGGCTGACCTGCCCCAGTTTATCTGGGGGATGAGAAAGATCTAGCGTGTGT	741
Db	770	TTTCGAGGCTTGAACTTACCACAATTCGAGATTCAGAAATATAAATAGCTCAATC--TGC	826
Oy	742	ACCAAGCACTACAAACACAGGGAATTCACCTGCATCGAGTAAAGTTTCACCTGAGACGG	801
Db	827	ATGGAAGAAATTTACATGCGAGAGTACAGCTGCCGGAAGGCCGACTTCACACTTGCAGGGG	886
Oy	802	CAGATGGGCTACTATCTGATTTGAGATGATACATCCCAAGCTAATCATGCTATCTGTCC	861
Db	887	TCAGTGGGCTACCAATGATGTCAGTGCATCTGCTACAGATGCTCATGCTGATCTG	946
Oy	862	TGGGTCCTCTTGATATCAATGATGATCTGCCCTGCCCCGTGTGGGCTGGGCAATCAC	921
Db	947	TGGGTCTCTTCTGCTGACGTGATGATGCATTTCCGGGGCCACACATGGGCTGCACG	1006
Oy	922	ACCGTGTACCATGATGACCAACCAGAGCTGTGGCTCCCGGCGCTTTGACTTAAGGTGTC	981
Db	1007	ACCGTGTCTACTATTTCTTCCAAAGGCGTCCGGTATACAGTCCCACTTGCTCCGGCTCA	1066
Oy	982	TACGTGAAGCGAATGACATCTGGATGGCTGTGTCTGTCTTTGTGTGCTGCTGCTTG	1041
Db	1067	TACGTGAAGGGAATCGATGTGTGATGAGAGCCGTGCACGGGCTTGTGTTCTCGGCACTA	1128
Oy	1042	CTGGAGTATGCTGCCATTA	1060
Db	1127	CTGGAGTTACCCGTGCTCA	1145
RESULT 14			
US-09-808-483-11			
Sequence 11, Application US/09808483			
Patent No. US2002001824A1			
GENERAL INFORMATION:			
APPLICANT: Bayer Aktiengesellschaft			
TITLE OF INVENTION: Ligand-gated anion channels of insects			
FILE REFERENCE: Le A 34 397			
CURRENT FILING DATE: 2001-03-14			
PRIOR APPLICATION NUMBER: US/09/808,483			
PRIOR FILING DATE: 2000-03-18			
NUMBER OF SEQ ID NOS: 18			
SOFTWARE: PatentIn Ver. 2.1			
SEQ ID NO 11			
LENGTH: 1458			
TYPE: DNA			
ORGANISM: Drosophila melanogaster			
FEATURE:			
NAME/KEY: CDS			
LOCATION: (1)..(1455)			
US-09-808-483-11			

Query Match	9.9%	Score 162.6;	DB 10;	Length 1458;
Best Local Similarity	52.3%	Pred. No.	3.9e-42;	
Matches	471:	Conservative	0:	Mismatches 379; Indels 51; Gaps 3
Oy	170	GATATGATGCCAGCATTTGGCCCAATTTTAAAGGCCACCGGTGAACGTACCTGCACA	229	
Dd	116	GTTACGACAMGATGAGACCGCGAAGAAGGGTCACCGCAGTAAGTACTTCCATG	175	
Oy	230	TCTTCATCACAGTTTGCTGCTCGTCACCACGACCAACGTAATGAGTCGGGTGAATGACT	289	

D	b		176	TGACGGTATGAGGTCTGGACTTCACATTGATGAAGAACTCGATGACTTATGTGGCGGATGCT	235
O	y	290	TCTTGCCGGCAACAGTAGAATAACCACACAGCCTGTCTCTAACCGAGATATC --- CTGATGACT	346	
D	b	236	TCTTTGCACAGACGtGGAGAAGATCATGCGCTCGGATTTGCCGAGAAATATGACNACGGAAT	295	
O	y	347	CTCTGGACCTGATGCCCTCCATCGCTGGACCTCATCTGAGAGCCAGACCTCTTTCTTGCTA	406	
D	b	296	AACCGCTCTCGAGAGTGAGTGGCTAAAAAAAATATGTGGCGGCCGGAATTCGTTTTTCAAA	355	
O	y	407	ATGAGAAAGGGGCCAATCTTCATGAGGAGACACGAGCAACAACATTAOTGGGACTCTTCA	466	
D	b	356	ACGCCAAATCCGTGACCTTTAGACACATGACAAATACCAAATCACTATATGTGGCTGTACA	415	
O	y	467	AGAATGGGAATGTGCTGTACAGACATCAAGGCTGACCCCTCATTTTGTCCCTGAGTGGACC	526	
D	b	416	AGGATAAACACATTTCTGTACATGGTCAAACTAACACATGAAAGCTTCTCGATCATGAACT	475	
O	y	527	TCAGAACTTCCCCATGAGACATCCAAGACGTGCACGATGACGCTTGAGAGCTCATCCATAC	586	
D	b	476	TCGCGATTATCCATCATGATACACAGGAGTGCACAGCTCAAGTCAAAATGAAAGCTGTTC	530	
O	y	587	TCTGAGCGCTCTGGCATCTCTGTACATTTGAGTTGGCTACACATGAAAGACCTGGT	646	
D	b	531	-----CCACACACGAGTACTTGATAT	553	
O	y	647	TTGAGTGGCTGGAAGATGCTCCTGCTGTCCAAGTGGCTGAGGGGCTGACTTGGCCAGT	706	
D	b	554	TCCAACTG -----GATCCAAACAGGCCCTTGTTGGTTGATGAAGAACATGGAATGCGCGC	607	
O	y	707	TTATCTTGGGGATGAGAAAGATCATAGGCTTTTATACCAAGCACTACAACACAGGAAT	766	
D	b	608	AGGTGGCCCTATCCGAATAAAGAGGGGACCTGTACCCAACTTATTTACCTAGTGGCACT	667	
O	y	767	TCACCTGATCGAGSTAAGTTTTCACCTGGAACGGACAGATGGGCTACTATCTGATTTAGA	826	
D	b	668	TCACATGGCTGSAAGTGGTGTTCACOCCTTAAGCCGCTTTGGTTACTAGCTTTCAACA	727	
O	y	827	TGTACATCCCACAGCTTACTCATTCGTATCTGTCTCTGGGTCTCCTTGTGATCAATGG	886	
D	b	728	CCTACATTTCCACACCTGCATGTAGTAGATCATGTATGCTTCCCTTGGATCAAAACGG	787	
O	y	887	ATGCGGCCCGCCCGGTGGGGCCGGGACATCACACCGCTCCATCCATGACCAACCCAGA	946	
D	b	788	AGGGCGCACACGCCGtGTGACTCTGGGTGTACACTCTTGTCTAAGCTTTCACGCAAC	847	
O	y	947	GCTCGGCTCCCGGGCCTCTTGTGCTAAGGTGTCTCAAGTGAAGCAATGCACATTCGA	1006	
D	b	848	AOGCAATTCGACAGTCTTTTGGCAACCTGTTTCTATCTCAADGCAAGTSGACCCCTTAA	907	
O	y	1007	TGGCTGTGTCTGCTCTTTTGTGTTCGCTGCTTGTCTGGAGTATGCTCATAAATTTG	1066	
D	b	908	TGTCCGTTTGCACGGTGTCTGTTTATGAGCCCTCATAGAGTATTTGTCTATAAATCATG	967	
O	y	1067	T. 1067		
D	b	968	T 968		

```

; RESULT 15
; US-09-808-483-9
; Sequence 9, Application US/09808483
; Patent No. US20020001824A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Aktiengesellschaft
; TITLE OF INVENTION: Ligand-gated anion channels of insects
; FILE REFERENCE: Le A 34 397
; CURRENT APPLICATION NUMBER: US/09/808.483
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: DE 100 136 19.2
; PRIOR FILING DATE: 2000-03-18
; NUMBER OF SEQ ID NOS: 18

```



SOFTWARE: PatentIn Ver. 2.1  
Seq ID NO: 9  
LENGTH: 1608  
TYPE: DNA  
ORGANISM: Drosophila melanogaster  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(1605)  
US-09-808-483-9

Query Match 9.9%; Score 162.6; DB 10; Length 1608;  
Best Local Similarity 52.3%; Freq. No. 4.1e-42;  
Matches 471; Conservative 0; Mismatches 379; Indels 51; Gaps 3;

```
OY 170 GATATGATCCAGGATTCGGCCCAATTTTAAAGGCCACCGTGAACGTGACCTGCACACA 229
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 116 GTTAGACAAATAGATGAGACCGCCGGAAGAGAGGTCACGACGATGACTTACTTCATG 175
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 230 TCTTCATCAACAGTTTCAGCTCCCTCACCAGACCCACATGAGTACCGGGTGAATGTCT 289
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 176 TGACGGTGTGATGGTCTGCTGACTCCATTGATGAGAACTGATGACTATGTGGCGATGTGT 235
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 290 TCTTGCGGCAACAGTGGGAAATGACCCAGCGCTGCTCTACCGAGAAATATC---CTGATGACT 346
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 236 TCTTTGCACAGACGTGGAGAGATCATCGCTGCGATGCGGAGAAATATGACACAGGAAT 295
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 347 CTCTGACCTCGATCCCTCCATGCTGACCTATCTGAGAGCCAGACCTCTTCTTGGCTA 406
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 296 ACCGCGCTCTGAGGTGACTGGCTAAATAATATGTGGCGCGGATTCGTTTTCAAAA 355
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 407 ATGAGAAAGGGGCCACTTCCATGAGTGCACGACGAGCAACAAGTTACTGCGCATCTTCA 466
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 356 ACGCAAAATCGGTGACTTTGACACATGACAATACCAATCACTATATGTGGCTGTACA 415
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 467 AGAATGGGAATGTGCTGACAGCATCAGGCTGACCCCTATTTGTCTGCTGATGAGACC 526
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 416 AGGATAAGACCATTTCTGTACATGGTCAACATGACATGACCTGCTGACATGAACT 475
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 527 TCAAGAACTTCCCATGAGATCCAGATCCAGATCAGATCAGCTTGAAGATCATCATATAC 586
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 476 TCGCATTTATCTCATGATATACAGAGTGCAGAGCTGCAATGGAAGCCCTGTC----- 530
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 587 TCTGACGCCCTCTGCGCATCTGTCTCATTTGCTTACAGCCATGGAAGACCTGCTGT 646
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 531 -----CCACACCAAGGATGACTGTATAT 553
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 647 TTGATGCTGGAAGATGCTCTGCTGTCGAAGTGGGTGAGGGGCTGACTCTGCCCACT 706
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 554 TCCAGTGG-----GATCCAACAACGCCCTTGTGTTGATGAAAAACATCGAATGCCGC 607
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 707 TTATCTTCGGGATGAGAAGATCTAGGCTGTGTGACCAAGCACTACACAACAGGGAAT 766
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 608 AGGTGGCCCTCATCCGAATATGAACGGGGAAGTGTACCAAGTTATTCACATGCAACT 667
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 767 TCACCTGATGAGGTAAAGTTTACCTGGAACGAGATGGGCTACTATCTGATTCAGA 826
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 668 TCACATGCTGAGGTGCTGTACACCTTAAAGCGTCTTGTGTTACTACGTTTCAACA 727
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 827 TGTACATCCAGGCTTACTCATGTCATGTCGTCTGCTGGTCTCTTGTGATCAACATGG 886
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 728 CCTACATTTCCACCTGCATGTAGTATGATCATGTATGGGTATCTTGTGATCAACCGG 787
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 887 ATGCTGCCCTTGCCGCTGTGGGCTGGCATCACACCGTGTACACATGACACCAACAGA 946
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 788 AGGGGGACACAGCCGCTGTGACTGTGGTGTACACTCTTGTGTACAGCTTTCCACGAC 847
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 947 GCTTGCTCCCGGCTCTTTTGGCTTAAAGTGTCTACGTGAAGCAATCGACATGTGA 1006
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 848 ACGCAAAATCCAGTCTTGTGGCACCTGTTTCTATCTCAAGGACAGTGGACGCTTTA 907
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1007 TGGCTGTGTCTGCTTTTGTGCTGCTGCTGCTGAGATGCTGCCAATAATTTTG 1066
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 908 TGTCCGTTTGACGGGTTCGTGTTTATGAGCCCTCATGAGATATGTCTAATAAACATCG 967
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

OY 1067 T 1067  
Db 968 T 968

Search completed: June 30, 2003, 23:06:59  
Job time : 269 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2003, 20:56:06 ; Search time 2191 Seconds  
(without alignments)  
12122.592 Million cell updates/sec

Title: US-10-075-846-3  
Perfect score: 1640  
Sequence: 1 atgcaactctgtctctgc.....taaaaaaaaaaaaaaaaaa 1640

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: em\_estlba:\*  
2: em\_estlhm:\*  
3: em\_estln:\*  
4: em\_estmu:\*  
5: em\_estrov:\*  
6: em\_estrpl:\*  
7: em\_estro:\*  
8: em\_estr:\*  
9: gb\_estl:\*  
10: gb\_est2:\*  
11: gb\_est3:\*  
12: gb\_est4:\*  
13: gb\_est5:\*  
14: em\_estfun:\*  
15: em\_esthm:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	401.8	24.5	662	9	AU169868 AU169868
2	302.4	18.4	817	12	BG404477 BG404477
3	295.6	18.0	678	10	AV729257 AV729257
4	292.4	17.8	779	14	B0443713 B0443713
5	266.6	16.3	692	12	BE981841 BE981841
6	255.2	15.6	720	12	BE981459 BE981459

7	239.6	14.6	552	13	BM440664	BM440664	pr.in.pX0
8	219.4	13.4	2991	11	BC027094	BC027094	Mus muscu
9	219	13.4	2593	11	BC022502	BC022502	Homo sapi
10	215.6	13.1	536	12	BF906462	BF906462	IL0-OT012
11	197.6	12.0	762	13	B1736322	B1736322	603360026
12	173.2	10.6	893	9	AL538200	AL538200	AL538200
13	171.2	10.4	908	13	B1825190	B1825190	603072067
14	168.6	10.3	501	10	AW280976	AW280976	f148d02.y
15	159.6	9.7	975	14	B0938794	B0938794	AGENCOURT
16	154	9.4	2475	11	AK013727	AK013727	Mus muscu
17	151	9.2	867	12	BC188196	BC188196	RST8236.A
18	147.4	9.0	331	12	BE830621	BE830621	PM2-MT002
19	145.4	8.9	817	9	AU133223	AU133223	AU133223
20	145.2	8.9	1942	11	BC029850	BC029850	Homo sapi
21	144.2	8.8	663	10	BB653397	BB653397	BB653397
22	143.6	8.8	690	14	B0044592	B0044592	UT-M-EO0
23	143.6	8.8	802	17	CNS02366Q	CNS02366Q	AL229643
24	143.6	8.8	1043	14	B0067799	B0067799	AGENCOURT
25	140.6	8.6	254	12	BC186166	BC186166	RST5011.A
26	138.4	8.4	583	12	BG203115	BG203115	RST22484
27	138.4	8.4	592	12	BG193387	BG193387	RST12516
28	137.8	8.4	317	12	BE830616	BE830616	PM2-MT002
29	137.8	8.4	804	12	BC196871	BC196871	RST16098
30	137.8	8.4	843	13	B1913857	B1913857	603183256
31	137.8	8.4	349	14	F06915	F06915	HSC10G061.n
32	136.8	8.3	567	13	BM129587	BM129587	1f20e03.y
33	135.8	8.3	570	13	BM129116	BM129116	1f17g06.y
34	135.8	8.3	723	13	BI736340	BI736340	603360049
35	135.2	8.2	792	12	BI736340	BI736340	RST5010.A
36	135.2	8.2	796	17	CNS0303VP	CNS0303VP	AL247390
37	135	8.2	975	13	BI552703	BI552703	603193953
38	134.6	8.2	785	14	BM964113	BM964113	UT-M-EO0
39	134.4	8.2	554	13	BM587895	BM587895	170006873
40	134.2	8.2	861	13	BI535004	BI535004	603193642
41	133.2	8.1	791	12	BG404434	BG404434	602420552
42	132.6	8.1	308	14	F13483	F13483	HSC3D6111.n
43	132.6	8.1	465	13	BI392407	BI392407	P991n.pX0
44	131.6	8.0	450	14	R13309	R13309	yf25d07.r1
45	131.2	8.0					

## ALIGNMENTS

RESULT 1  
LOCUS AU169868 662 bp mRNA linear EST 29-JAN-2001  
DEFINITION AU169868 Ol-br-ad cDNA Oryzias latipes cDNA clone br5332, mRNA  
ACCESSION AU169868  
VERSION AU169868.1 GI:12591937  
KEYWORDS EST.  
SOURCE Japanese medaka.  
ORGANISM Oryzias latipes  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
Acanthomorpha; Acanthopterygii; Percomorpha; Alburnomorpha;  
Belontiiformes; Adiantichthyidae; Oryziinae; Oryzias.  
REFERENCE 1 (bases 1 to 662)  
Mita, K., Ishikawa, Y. and Yamauchi, M.  
Establishment of cDNA database of medaka, Oryzias latipes  
Unpublished (2001)  
COMMENT Contact: Mita K  
Genome Research Group  
National Institute of Radiological Sciences  
Anagawa 4-9-1, Inage, Chiba 263-8555, Japan  
Email: kmila@nirs.go.jp  
method: uni-directional sequence direction: sequenced from T3 primer  
(5' -> 3')  
FEATURES  
source location/Qualifiers  
1..662  
/organism="Oryzias latipes"  
/strain="HNI"

```

BASE COUNT      155 a      205 c      164 g      138 t
ORIGIN

Query Match      24.5% Score 401.8; DB 9; Length 662;
Best Local Similarity 76.4%; Pred. No. 2e-74; Indels 45; Gaps 2;
Matches 540; Conservative 0; Mismatches 122;

OY 195 TTTTAAAGGCCCCCGTGAAGCTGACCTGCAACATCTTCATCAACAGTTTCAGTCCGT 254
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 TTTTAAAGTCCACCTGTAAGCTGACCTGCAACATTTTATCAACAGTTTGGATTCAT 60

OY 255 CACCAAGACCCACATGACATGACCGGTGTAATGCTTTCTGGCGCAACAGTGAATGACCC 314
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 CGCTGAACACACCATGACATGACGAGTGAACATCTTCGAGGACAGAGTGAACGACCC 120

OY 315 AGCGCTGCTACGAGATATCGATGACATCTGACCTGACCTGATCCCTCATGCTGGA 374
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 121 CGGCTGGCTACGACGAGATATCCGACGACTGCTGATCTGACCCCTCATGTTGA 180

OY 375 CTCTATCTGGAAGCCAGACCTCTCTTTCCTAATGAGAAAGGGCCAACTTCATGAGCT 434
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 181 CTGCACTGGAAGCCGAGATCTGCTTCTTCTAATGAGAAAGGGCCAACTTCACAGAGT 240

OY 435 GACCAAGCAACAAGTCTGCGCATCTCAAGAATGGGAATGCTCTACAGATCAG 494
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 241 CACCAAGCAACAAGTCTGCGCATCTCAAAAATGGCAATGCTCTACAGATCAG 300

OY 495 GCTGACCCATCTTGTCTGCTGCTGAGACCTCAAGACTTCCCATGAGATCCAGAC 554
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 301 AATCACTGCTGATCTGCGCTGCGCCATGATCTGAAGACTTCCCATGAGATGTCAGAC 360

OY 555 GTGACGATGACAGCTTGAAGCTCATCTACTCTGACGCTCTGCCATCTCTGTCAT 614
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 361 CTGCACTCAACAGCTGAGAGCT----- 383

OY 615 TTCAGTTGGCTACACCATGAAGAGCTGCTGTTGAGTGGCTGAGAGATGCTCTGCTGT 674
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 384 -----TCGGCTACACCATGAAGAGCTCATCTTCGAGTG-----GACAGAAAGGGCGCT 435

OY 675 CCAAGTGGCTGAGGGGCTGCTGCTGCGCAGTTTATCTGCGGATGAGAGAGATCTAGG 734
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 436 GCAAGTGGCGCAAGGGCTGACGCTGCTCAGTTCATCTGAAGAGAGAGAGAGCTGCG 495

OY 735 CTGTTTACCAAGCACTACACAAGGGAATTCACCTGATGAGAGTAAAGTTTACCT 794
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 496 CTACTCTACCAAGCACTACACAAGGTAATTCACCTGATGAGAGTAAAGTTTACCT 555

OY 795 GGAAGGGAGATGGGCTGATCTATGATGATGATGATGATGATGATGATGATGATGAT 854
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 556 GGGGCGACAGATGGGCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 615

OY 855 CCTGTCTCTGGTCTCTCTCTGATCAACATGATGATGATGATGATGATGATGATGAT 901
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 616 CCGTCTCTGGTCTCTCTCTGATCAACATGATGATGATGATGATGATGATGATGAT 662

RESULT 2
LOCUS      BG404477      817 bp      mRNA      linear      EST 12-MAR-2001
DEFINITION 602420604F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:4527768 5',
ACCESSION  BG404477
VERSION    BG404477.1  GI:13297925
KEYWORDS   EST.
SOURCE     house mouse.
ORGANISM   Mus musculus.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

REFERENCE      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS      1 (bases 1 to 817)
TITLE      NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL      National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT      Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cga@bgl-remail.nih.gov
            Tissue Procurement: The Cepko Laboratory
            cDNA Library Preparation: Life Technologies, Inc.
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNLN at:
            http://image.llnl.gov
            Plate: LLM10437 row: 9 column: 01
            High quality sequence start: 6
            High quality sequence stop: 808.

FEATURES
    source
        1..817
        /organism="Mus musculus"
        /db_xref="taxon:10090"
        /clone="IMAGE:4527768"
        /clone_lib="NIH_MGC_94"
        /tissue_type="retina"
        /lab_host="DH10B (phage-resistant)"
        /note="Organ: eye; Vector: pCMV-Sport6; Site:1: NotI;
        Site:2: SalI; Cloned unidirectionally; oligo-dT primed.
        Average insert size 3.3 kb. Library enriched for
        full-length clones and constructed by Life Technologies.
        Note: this is a NIH-MGC Library."

BASE COUNT      223 a      190 c      192 g      212 t
ORIGIN

Query Match      18.4% Score 302.4; DB 12; Length 817;
Best Local Similarity 71.6%; Pred. No. 1.5e-53;
Matches 496; Conservative 0; Mismatches 180; Indels 17; Gaps 7;

OY 697 CTGCCCCATTTATCTGCGGATGAGAGATGAGTCTAGGCTGTTGACCAAGACTACAAAC 756
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 113 CTGCCCCATTTATTTTGAAGAGAGAGAGCTGTTGTTTGCACAAAGACTTACAAAC 172

OY 757 ACAGGAAATTCACCTGATTCAGAGTAAAGTTTCACTGAGGAGGAGGAGTGGCTACTAT 816
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 173 ACTGGCAATTTACCGTATGAGTGCAGTTTCACTGAGGAGGAGTGGCTACTAT 232

OY 817 CTGATTCAGATGATACATCCCAAGCTTCACTGATGATGATGATGATGATGATGATGAT 876
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 233 TTGATTCAGATGATACATCCCAAGCTTCACTGATGATGATGATGATGATGATGATGAT 291

OY 877 ATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 936
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 292 ATTAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 350

OY 937 ACCACCAAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 996
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 351 ACTACACAGATTCAGG-TTCCAGGATCTGCGCAAGGTCCTCATGTAAGCAAT 409

OY 997 GACATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1056
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 410 GACATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 469

OY 1057 ATTAATTTGTTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1116
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 470 GTGAA-CTTGTTCTCCAGGCACTAAGAGAGTTCTTCCTCCGAGAGAGAGAGAGAG 528

OY 1117 CAAGCTTGAAGAGATATCATCAAGAAAGTCTGTTTATTTCCGTTGATGATGATGATGAT 1176
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 529 CAGATTAAGAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 588

OY 1177 GGCACATGCTGCGAGGCAAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1236
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 589 GTTCACTGCTCCCAATGAAAGATG-----GCACAGCTGTCAAGGCTACACCTGC 638

```

Accession	Sequence	Position
OY	1237 CACCTTCGAGCCCTCTCTTAAGGGAAGAAACACACGCGAATCACTACGTGAGCTGA	1296
Db	639 CACCCCACTTCCACAAACCCCAAGAGTGCACATGCTATCAAGAAAGAGT-TGTGGATCGG	697
OY	1297 GCCAAGAGAAATGACACCATCTCCCGGGCTGTCTCCCTTCACTTTCCTCATCTCAAT	1356
Db	698 GCAAAAAGAAATGACACCATATCTCGAGACGTCTTCCCACT--GGCTTCCTATTATTCAAC	755
OY	1357 ATCTTCTACTGGTGTGCTATPAAAGTGCATANG 1389	
Db	756 ATCTTTACTGTGATCACAATACAAATCATTTCCG 788	

[illegible]

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

1 (bases 1 to 678)  
Gu, Y., Peng, Y., Song, H., Huang, Q., Yang, Y., Gao, G., Xiao, H., Xu, X.,  
Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng, L., Xu,  
S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Hu, R.,  
Chen, J., Chen, Z., and Han, Z.,  
Homo sapiens cDNA HTC clones  
Unpublished (2000)  
Contact: Zeguang Han

TITLE	JOURNAL	COMMENT
Homo sapiens CDNA HTC clones Unpublished (2000)		Contact: Zeguang Han Chinese National Human Genome Center at Shanghai 351 Guo Shouling Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China Tel: 86-21-50801919(ex.45) Fax: 86-21-50801922 Email: hanzg@chgc.sh.cn This clone is available at CHGC in Shanghai.

FEATURES	SOURCE
Location/Qualifiers	1. .678
/organism="Homo sapiens"	
/db_xref="taxon:9606"	
/clone="HMCHE03"	
/clone_lib="HRC"	
/tissue_type="Hypothalamus"	
/dev_stage="Adult"	
/lab_host="SOLR"	
/note="Vector: plasmidscript sk(-); Site_1: EcoRI; Site_2: XhoI"	
BASE COUNT	199 a 148 c 150 g 160 t
ORIGIN	1 others

	Query Match	18.0%	Score 295.6	DB 10	Length 678
	Best Local Similarity	74.9%	Pred. No. 4.3e-52		
	Matches	370	Conservative	0	Mismatches 12
				Indels	0
				Gaps	0
OY	84	AAAGAGAGCAATCTGGACCAAGGGTCCAGGCCATGTCCTCTGATTTCT	143		
Db	49	AGACCATGACTCACAGTGTGGAAAAACCACTTCACAGACCCTATCTCTCAGATTTCTT	108		
OY	144	AGACCAACTTATTGGGGGCACATCGATGATGATGATGCCAGTMTGGCCCATTTTAAAG	203		
Db	109	GGACAAGTTAATGGAMGACATCGATGATGATGATGCAAGATTCAGGCCAAATTTTAAAG	168		
OY	204	CCCAACCCGTGACGTGACCTGCAACATCTTCATCAACAGTTTCAGCTTCCTGACCAAGAC	263		
Db	169	TCCTCCATTAACGTTACTTCCATATTTTATCAACAGTTTGGATCAGTCACAGAAC	228		
OY	264	CACATGACATCCGGGTGATGTCCTTTGCGGCAAGATGGAATGACCAAGCCTGTC	323		
Db	229	GACCATGGACATCCAGATGATATTTTCTAGACAACAGTGGATGATTCACGGCTGGC	288		

QY	324	TTACGGAATATCTGTGACCTGTGGACCTGACCTCCCAATGGTGGACTATCTG	383
Db	289	GTACAGTAGTACCAATGACTCCCTGGACTTGGACCATCATGCTAGACTCATTTTG	348
QY	384	GAAGCCAGACCTCTTCTTGTGTAATGAGAAAGGGCCAACTTCCATGAGGTGACCA	443
Db	349	GAACCAAGATTGTTCTTTGGCAATGAGAAAGGGTCCAACTTCCACGATGCACACTGA	408
QY	444	CAACAAGTACTGGGCATCTCAAGATGGAATGTGCTGTACAGCATTCAGGCTGACCCT	503
Db	409	CACAATAATTGCTACGCAATTTGAAAAATGGCAAAAGTGCTTACAGTATTCAGACTT	468
QY	504	CATTTTGCTCCCTGATGACACTCAAGAACTTCCCATGAGACATTCAGAGCTGACAGT	563
Db	469	GACCTTATCCGTGCCCATGACTGAGAAACTTTCGATGATGTCCAGACCTGTACAAT	528
QY	564	GCAGCTTGAGAGCT 577	
Db	529	GCAGCTGAGAGTT 542	

<b>RESULT 4</b>					
B0443713					
<b>LOCUS</b>	B0443713	779 bp	mRNA	linear	EST 29-MAY-2002
<b>DEFINITION</b>	U1-M-EW0-bxh-j-02-0-U.r1 NIH.BMAP_EW0 Mus musculus cDNA clone				
<b>IMAGE:</b>	5708257 5', mRNA sequence.				

ACCESSION	BC043713
VERSION	BC043713.1
KEYWORDS	GI:21246825
SOURCE	ESN.
ORGANISM	house mouse.
REFERENCE	Mus musculus.
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE	1. (bases 1 to 779)
JOURNAL	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
CONTACT	Contact: Robert Strausberg, Ph.D.

Email: c9apbs@email.nih.gov  
Tissue Procurement: Dr. James Lin, University of Iowa  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
This clone was contributed by the Brain Molecular Anatomy Project  
(BMAP)  
Seq primer: pYX-5.

```

FEATURES
source
Location/Qualifiers
1..779
/organism="Mus musculus"
/strain="C57BL/6"
/db_xref="taxon:10090"
/image="5708257"
/clone_lib="NIH_BMAP_EMO"
/tissue_type="whole brain"
/dev_stage="embryo 15.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/note="Organ: brain, Vector: pYX-Asc, Site_1:
Site_2: Not I; The library was constructed according to
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured mRNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was size selected according to mRNA size fraction,
ligated with EcoR I adaptor, digested with Not I, and then
cloned directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
, is GTGGCGGCAA. This library was created for the
University of Iowa Mouse Brain Molecular Anatomy Project
(BMAP). 'Gene Discovery in the Developing Mouse Nervous
System', supported by National Institutes of Mental Health

```

(NIMH), Hemin Chin, Ph.D., program coordinator."

BASE COUNT  
ORIGIN

222 a 167 c 165 g 225 t

Query Match 17.88; Score 292.4; DB 14; Length 779;  
Best Local Similarity 74.58; Pred. No. 2e-51;  
Matches 368; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

84 AAAAGAGAGTCAATCTGGAACCAAGGGGTCCGCCATGTCCTCTGATTCCT 143  
197 AGACATGACTCCAGCTGTGGAACATCCCTCGACACCTATCTCTTCACTTTCTT 256  
144 AGACAACTTATGAGGGGCAACATCTGATATGATGCGAGATTCGCCCAATTTAAAG 203  
257 GATTAATTAATGAGGAGCAATCATGATGATGATGATGATGATGATGATGATGATGAT 316  
204 CCCACCCGTAAGTGAATCTGCAACATCTTCAACATGATGATGATGATGATGATGATGAT 263  
317 TCTCCAGTAACTGATCTGCAATATTTTATCAAGTTTGGATGATGATGATGATGATGAT 376  
264 CACATGACTACCGGGTGAATGCTCTTCTGCGGCAAGTGAATGACCCAGCTGTC 323  
377 CACATGACTACCGAGTGAACATTTTCTGAGACAGTGAATGATGATGATGATGATGATGAT 436  
324 CTACCGGAATATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 383  
437 ATACAGTACCTACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 496  
384 GAAGCCAGCT 443  
497 GAACCGGATTTTCT 556  
444 CAACAAGTACTGCGATCTTCAAGATGGAATGCTGATGATGATGATGATGATGATGATGAT 503  
557 CACAAGTCTTGTGGATTTCCAAATGCAAGTCTCTACAGTATTTAGACTCAGCTT 616  
504 CATTTTCT 563  
617 GACTTATCTTGTTCATGACCTGGAAGTCTTCCATGATGATGATGATGATGATGATGATGAT 676  
564 GCAGCTGAGAGCT 577  
677 GCAGCTGAGAGCT 690

RESULT 5  
BE981841 692 bp mRNA linear EST 29-APR-2002  
LOCUS  
DEFINITION  
UT-M-CGDP-bdd-h-07-0-UI.s1 NIH\_BMAP\_Ret4\_S2 Mus musculus cDNA clone  
UT-M-CGDP-bdd-h-07-0-UI 3, mRNA sequence.

ACCESSION  
BE981841  
VERSION  
BE981841.1 GI:10651376  
KEYWORDS  
EST.  
SOURCE  
house mouse.  
ORGANISM  
Mus musculus.

REFERENCE  
AUTHORS  
TITLE  
Normalization and subtraction: two approaches to facilitate gene  
discovery

JOURNAL  
MEDLINE  
97044477  
Contact: Chin, H  
National Institute of Mental Health  
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD  
20892-9643, USA  
Tel: 301 443 1706  
Fax: 301 443 9890  
Email: meste@mail.nih.gov

oligo-ct track not found, Not I site shown in beginning of sequence  
is likely internal to the message. cDNA library preparation: M.B.  
Soares Lab Clone distribution: Researchers may obtain BMAP cDNA

clones from RESEARCH GENETICS. It should be noted that Bento Soares  
is generating a small number of additional specialized  
non-redundant arrays of BMAP cDNAs whose availability will be  
considered under appropriate and limited collaborative arrangements  
The tissue for this library was contributed by Dr. Xin-Yuan Fu,  
Yale University School of Medicine  
Seq primer: M13 Forward  
POLYA-No.

FEATURES  
source

location/Qualifiers  
1..692  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UI-M-CGDP-bdd-h-07-0-UI"  
/clone\_1lb="NIH\_BMAP\_Ret4\_S2"  
/lab\_host="DH10B (Life Technologies)"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker. Site 1: Not I; Site 2: Eco RI; The  
NIH\_BMAP\_Ret4\_S2 library is a subtracted library,  
ultimately derived from mouse retina tissue libraries at  
various stages of development. For a detailed description  
of the library from which this clone was derived, please  
visit our web site at brainest.eng.yuwa.edu. The tissue  
for this library was contributed by Dr. Xin-Yuan Fu, Yale  
University School of Medicine  
TAG\_SEQ=None found"

BASE COUNT 190 a 159 c 144 g 198 t 1 others  
ORIGIN

Query Match 16.38; Score 266.6; DB 12; Length 692;  
Best Local Similarity 70.98; Pred. No. 5.4e-46;  
Matches 353; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

80 TGGCAAGAGGAAAGTCAATCTGGAACCAAGGGTCCCGCATGTCCTCTCTGAT 139  
172 TGGCCACAAAGGAAACAAACAGTCCGATCTCGAATGCTCAATGTCCTCTCTGAT 231  
140 TCTAGCAAACTTATGGGCGGAACATCTGATATGATGCGAGATTCGCCCAATTTA 199  
232 TTCTGGAACAAATATGGGAGGAGCATGGGGATATGATGCAAGATCAAGCCACTTCA 291  
200 AAGGCCACCCGTAAGCTGACCTGACACATCTTCAATCAAGTTCAGCTCCGTACCA 259  
292 AAGTCTCTCCAGTAAATGTCACATGCAATATTCATTAACACTTGGCTCCATTGCA 351  
260 AGACCAATGAGTACCGGGTGAATGCTCTTGGCGCAACAGTGAATGACCCAGCC 319  
352 AGACGACTATGATTTACAGATTAACATTTTCTCTCGAAGTGAATGATGCTGCTC 411  
320 TGTCTACCGAGATATCTGATGACTCTGACCTGATCCCTCATGCTGGACTCTA 379  
412 TTGCATGACGATGATATCTGACGATCATTTAGACCTGACCGGTGATTTGGATTCA 471  
380 TCTGGAAGCCAGACTTTCTTTGCTATATAGAAAGGGCCCAACTTCATGAGTGACCA 439  
472 TTTGGAACCTGACTGTTTCTTCTATATAGAAAGGGGCTAATCCACAGAACTACCA 531  
440 CGGACCAAGTACTGCGCATCTCAAGATGGAATGCTGTACAGCATGAGGTGA 499  
532 CAGATTAACAGCTGCTAGATTTTTCANAAATGGAATGCTTTTATCAATAGGTGA 551  
500 CCTCATTTTCTCTGCTGCTGACCTCAAGACTTCCCATGACATGACAGCTGCA 559  
592 CATTAACAGCTCTCTGCAATGATGATCAAGATTTCCCAATGATGATCAACATGCA 651  
560 CGATGACGCTTGAGAGCT 577  
652 TAATGCACTTGAAAGCT 669

RESULT 6  
BE981459 720 bp mRNA linear EST 29-APR-2002  
LOCUS

DEFINITION UI-M-CG0P-bdc-d-08-0-UI-s1 NIH\_BMAP\_Ret4\_S2 Mus musculus cDNA clone  
 UI-M-CG0P-bdc-d-08-0-UI 3', mRNA sequence.  
 BE981459  
 ACCESSION BE981459.1 GI:10650587  
 VERSION EST.  
 KEYWORDS house mouse.  
 SOURCE Mus musculus  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerognathii; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 720)  
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.  
 TITLE Normalization and subtraction: two approaches to facilitate gene discovery  
 JOURNAL Genome Res. 6 (9), 791-806 (1996)  
 MEDLINE 97044477  
 COMMENT Contact: Chln. H  
 National Institute of Mental Health  
 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD  
 20892-9643, USA  
 Tel: 301 443 1706  
 Fax: 301 443 9890  
 Email: mestr@mail.nih.gov  
 Oligo-dT track not found, Not 1 site shown in beginning of sequence  
 is likely internal to the message. cDNA Library Preparation: M.B.  
 Soares Lab Clone distribution: Researchers may obtain BMAP cDNA  
 clones from RESEARCH GENETICS. It should be noted that Bento Soares  
 is generating a small number of additional specialized  
 non-redundant arrays of BMAP cDNAs whose availability will be  
 considered under appropriate and limited collaborative arrangements  
 The tissue for this library was contributed by Dr. Xin-Yuan Fu,  
 Yale University School of Medicine  
 Seq primer: M13 Forward  
 POLYA-No.

FEATURES  
 SOURCE location/Qualifiers  
 1..720  
 /organism="Mus musculus"  
 /strain="C57BL/6j"  
 /db\_xref="taxon:10090"  
 /clone="UI-M-CG0P-bdc-d-08-0-UI"  
 /clone.lib="NIH\_BMAP\_Ret4\_S2"  
 /lab\_host="DH10B (Life Technologies)"  
 /note="Vector: pT73D-Pac (Pharmacia) with a modified  
 polylinker; Site\_1: Not I; Site\_2: Eco RI; The  
 NIH\_BMAP\_Ret4\_S2 library is a subtracted library,  
 ultimately derived from mouse retina tissue libraries at  
 various stages of development. For a detailed description  
 of the library from which this clone was derived, please  
 visit our web site at brainest.eng.uiowa.edu. The tissue  
 for this library was contributed by Dr. Xin-Yuan Fu, Yale  
 University School of Medicine  
 TAG\_SEQ=None found"

BASE COUNT 198 a 162 c 206 t 3 others  
 ORIGIN

Query Match 15.6%; Score 255.2; DB 12; Length 720;  
 Best Local Similarity 70.7%; Pred. No. 1.3e-43;  
 Matches 352; Conservative 0; Mismatches 145; Indels 1; Gaps 1;

QY 80 TGGCAAAGAGAGTCAATCTGGAACCAAGGGCTCCAGCCCATGTCCTTGAT 139  
 DB 171 TGGCCCAAGAAACAAACAGATCGCGATCTCGAAGTCTTCATGTCACCTTCTGATT 230  
 QY 140 TCTTGACAACATTTATGGGCGAAGCATCTGATATGATGCGAGATTGGCCCAATTTTA 199  
 DB 231 TTCTTGACAACATTTATGGGCGAAGCATCTGATATGATGCGAGATTGGCCCAATTTCA 290  
 QY 200 AAGGCCACCCGTGAGCTGACCTGCAACATCTTCATCAACAGTTTCAGCTCCGTACCA 259  
 DB 291 AAGGCTCTTCAGATTATGTCACATGCAACATATTCATAACAGCTTTGGCTCCATTGCGAG 350  
 QY 260 AGACACATGAGACAGCGGGTGATGCTTCTTGGCGCAACAGAGATGACCCACGCC 319  
 DB 351 AGACAGTATGAGATTACAGATTAAACATTTTCTTCTGACAGAGATGATGATCTCGTC 410

QY 320 TGTCTACCGAGATATCTGATGACTCTGTGACCTCGATCCCTCCAGTGCATCTGA 379  
 DB 411 TTGCATACAGTGAATATCTCTGACGATTTATTAGACTTGACCCGCTATGTTGGATTCCA 470  
 QY 380 TCTGGAAGCCAGACCTCTTCTTGTCTAATGAGAAAGGGGCCAACTTCATGAGTGACCA 439  
 DB 471 TTTGGAAACCTGACTGTTGTTCTTCTAATGAGAAAGGGGCTTAACCTTCACGAGTACCA 530  
 QY 440 CGGACAACAGTACTGCGCATCTTCAAGAAATGGAATGTGCTGACAGCATGAGCTGA 499  
 DB 531 CAGATTAACAGCTGCTAAGAAATTTTCANAAAGAAATGCTTTATTCATTAAGGTTGA 590  
 QY 500 CCCGATTTTTCGCTGCGCTGATGACCTCAAGAACTTCCCATGGACATCCAGAGCTCA 559  
 DB 591 CATTAACACTCTCCCTGCTCCATGATGATCTCAGAA-TTTCCAAATGATGATACACATGCA 649  
 QY 560 CGATCGACCTTGAGAGCT 577  
 DB 650 TAATGCAACTTGAAGCT 667

RESULT 7  
 BM440664  
 LOCUS  
 DEFINITION 552 bp mRNA linear EST 01-FEB-2002  
 (pgrln.pk003.m19 Normalized Chicken Reproductive Tract cDNA Library  
 (pgrln).Gallus gallus cDNA clone pgrln.pk003.m19 5' similar to  
 p11M49701M9970 glycine receptor alpha-4 chain - mouse (fragment  
 ), mRNA sequence.  
 BM440664  
 ACCESSION BM440664.1 GI:18471439  
 VERSION EST.  
 KEYWORDS chicken.  
 SOURCE  
 ORGANISM Gallus gallus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Archosauromia; Aves; Neognathae; Galliformes; Phasianidae;  
 Phasianinae; Gallus.  
 REFERENCE 1 (bases 1 to 552)  
 COGBURN,L.A. and NYS,Y.  
 ESTs from Normalized Chicken Reproductive Tract cDNA Library-  
 University of Delaware and INRA, Tours-Poultry Unit Project  
 Unpublished (2002)  
 JOURNAL Contact: Larry A. Cogburn  
 University of Delaware  
 Townsend Hall, Newark, DE 19717, USA  
 Tel: 302-831-1335  
 Fax: 302-831-2822  
 Email: cogburn@udel.edu, www.chickest.udel.edu.

FEATURES  
 source location/Qualifiers  
 1..552  
 /organism="Gallus gallus"  
 /strain="Commercial broiler and layer"  
 /db\_xref="taxon:9031"  
 /clone="pgrln.pk003.m19"  
 /clone.lib="Normalized Chicken Reproductive Tract cDNA  
 library (pgrln)"  
 /sex="Male and Female"  
 /tissue\_type="testis, ovary and oviduct"  
 /dev\_stage="Various stages: embryonic, post-hatch, immature  
 and sexually-mature"  
 /lab\_host="E. coli EMDH10B"  
 /note="Vector: PCMVSPORT6; Library made from three total  
 RNA pools from each tissue (testis 25%, ovary 25%, and  
 oviduct 50% of final RNA pool); Single pass sequencing  
 from 5'-end"

BASE COUNT 85 a 198 c 181 g 88 t  
 ORIGIN

Query Match 14.6%; Score 239.6; DB 13; Length 552;  
 Best Local Similarity 76.9%; Pred. No. 2.7e-40;  
 Matches 306; Conservative 0; Mismatches 89; Indels 3; Gaps 1;

QY 19 GCACCCCTCTCTCTTCTTCTGAGACCCCTGCAGGGCAGGCTCTCTCAGGCTGCC 78









```

SOURCE      human.
ORGANISM    Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE   1 (bases 1 to 893)
AUTHORS    Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE      Full-length cDNA libraries and normalization
JOURNAL    Unpublished (2001)
COMMENT     Contact: Genoscope
             Genoscope - Centre National de Sequencage
             BP 191 91006 Evry cedex - France
             Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
    source
        1..893
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="CSDF029P18"
        /clone_1lb="LTI_FL013_Fbrn1"
        /dev_stage="pooled tissue from post conception fetuses (20
        week, 24 week and 26 week)"
        /lab_host="DH10B"
        /note="Organ: Fetal brain; Vector: pCMVSPORT 6; 1st strand
        cDNA was primed with a NotI-oligo(dT) primer. Five prime
        end enriched, double-stranded cDNA was digested with Not I
        and cloned into the Not I and Eco RV sites of the
        pCMVSPORT 6 vector. Library was constructed by Life
        Technologies. Contact : Feng Liang Life Technologies, a
        division of Invitrogen 9800 Medical Center Drive Rockville
        , Maryland 20850, USA Fax : (1) 301 610 8371 Email :
        fliang@life.com URL :
        http://fulllength.invitrogen.com"
BASE COUNT  235 a      188 c      209 g      253 t      8 others
ORIGIN
Query Match      10.6% Score 173.2; DB 9; Length 893;
Best Local Similarity 54.6% Pred. No. 2.3e-26;
Matches 389; Conservative 5; Mismatches 276; Indels 42; Gaps 1;

OY 343 GACCTCTGAGACCTGATCCCTGATGAGCTGATCTATCTGAGAGCCGACCTCTCTT 402
    || || || || || || || || || || || || || || || || || || || || ||
DB 54 GATGACGACAGCTGGATGACCAACATGTTATGAAACCTGATTTATTTT 113
OY 403 GCTATGAGAAAGGGCCCACTCCATGAGGTGACACGACACAACTACTGCGCATC 462
    || || || || || || || || || || || || || || || || || || || || ||
DB 114 GCAATGAAAAAAGGCCAATTTTCATGATGACCCAGAAAACATCTCTCTTTAT 173
OY 463 TTCAAGATGGGAATGCTCTACAGCATCAGGCTGACCTCATTTTCTCTCGTATG 522
    || || || || || || || || || || || || || || || || || || || || ||
DB 174 TTTCGTATGAGATGCTCTGTCTGACATGAGTTATCTATTACTTTCAATGCCCTT 233
OY 523 GACCTCAAGACTTCCCATGAGATCCAGACGTGACAGCTGAGAGCTCATCC 582
    || || || || || || || || || || || || || || || || || || || || ||
DB 234 GACTTGACATTTGTTCCATGATGACACAGATGCAAGATGCACTGAGAGCT---- 288
OY 583 ATACTGACAGCCCTGTCATCTCTGCTCACTTTCATGCTACACCATGAAGACCTC 642
    || || || || || || || || || || || || || || || || || || || || ||
DB 289 -----TTGGTTACAAACATGATGATT 311
OY 643 GTGTTTGAGTGGCTGGAAGATGCTCTGCTCCAGTGGCTGAGGGCTGACTGCCC 702
    || || || || || || || || || || || || || || || || || || || || ||
DB 312 CGATTTATCTGGAGTCAGAGATCCCTGCAATTAATAAATTTGCCCTCAATTT 371
OY 703 CAGTTTATCTGGGATGAGAGATCTAGCTGTTTACCAAGCACTATACACAGCAGG 762
    || || || || || || || || || || || || || || || || || || || || ||
DB 372 GATATCAAAAAGGAAGATATGTAATGTAATCTAATAAATACTATAAAGCCAGCAG 431
OY 763 AAATTCACCTGCATGAGTAAGTTTCACTGGAACGAGAGTGGCTACTATCTGATT 822
    || || || || || || || || || || || || || || || || || || || || ||
DB 432 AACAAAMAAADBKTTGATGATCTTCAACCTGAGAGGAGCAGGTGGCTTTTACATG 491
OY 823 CAGATGTACATCCCGACGCTACTATCTGCTCTGCTGCTCTCTGATGATCAG 882
    || || || || || || || || || || || || || || || || || || || || ||
DB 492 GGGGTCTACGGCCCAACCTGCTCATTTGTTCTCTGCTGCTCTCTGATCAAC 551

```

```

OY 883 ATGATGTCGCCCGCCGCGCCGCTGGGCTGGGACATCAACACCGCTACCATGACCACC 942
    || || || || || || || || || || || || || || || || || || || || ||
DB 552 CCGAGCGGAGTGTCTGACAGATGCCCCCTGGATCTTCTAGTCTCAGCTTGCCCT 611
OY 943 CAGAGCTCTGCTCCCGGCCCTCTTTCCTTAAGGTGTCATGTAAGGCAATGACATC 1002
    || || || || || || || || || || || || || || || || || || || || ||
DB 612 GAGTGCAACACCTTTGGCCGCTGAGCTTCCCAAGTTTCTATGGAAGCCTTGATGTT 671
OY 1003 TGGATGCTGTGTCTGCTCTTCTTGTGTCCTGCTGCTGCTGAGATGCTG 1054
    || || || || || || || || || || || || || || || || || || || || ||
DB 672 TGGCTTATGCTGCTCTCTTCTTGGGCTTCTCCCGTGGAGATGACAG 723

RESULT 13
BI825190      908 bp mRNA linear EST 04-OCT-2001
LOCUS        603072067F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5164025 5',
DEFINITION   mRNA sequence.
ACCESSION    BI825190
VERSION      BI825190.1 GI:15936740
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE   1 (bases 1 to 908)
AUTHORS    NIH-MGC http://mgc.ncl.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
             Email: c9qab@femail.nih.gov
             Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
             found through the I.M.A.G.E. Consortium/LNL at:
             http://image.lnl.gov
             Plate: LNL1407 row: e column: 18
             High quality sequence stop: 844.

FEATURES
    source
        1..908
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="IMAGE:5164025"
        /clone_1lb="NIH_MGC_119"
        /tissue_type="medulla"
        /lab_host="DH10B"
        /note="Organ: brain; Vector: pCMV-SPORT6; Site.1: NotI;
        Site.2: EcoRV (destroyed); RNA source normal medulla from
        anonymous male age 27. Library is oligo-dT primed and
        directionally cloned (EcoRV site is destroyed upon
        cloning). Average insert size 1.3 kb. Insert size range
        0.9-3 kb. Library is normalized and enriched for
        full-length clones and was constructed by C. Gruber
        (Invitrogen). Research Genetics tracking code 013. Note:
        this is a NIH_MGC Library."
BASE COUNT  341 a      156 c      190 g      221 t
ORIGIN
Query Match      10.4% Score 171.2; DB 13; Length 908;
Best Local Similarity 74.4% Pred. No. 6e-26;
Matches 229; Conservative 0; Mismatches 78; Indels 1; Gaps 1;

OY 271 GACTACCGGGTGAATGTC-TTCTTGGCGCAACAGTGAATGACCCAGCGCTCTACCG 329
    || || || || || || || || || || || || || || || || || || || || ||
DB 42 GATTAACAGAGATATATCTTCTTGGCGTACAGAAATGATGATCCCGCTGCGTACAG 101
OY 330 AGAATATCTGATGACTCTGAGACCTGATCCCTCATGCTGAGACTATCTGGAACC 389
    || || || || || || || || || || || || || || || || || || || || ||
DB 102 TGAATATCTGACGACTCTTATAGACCTTGACCCCTCATGTTGGACTCATTTTGAAC 161

```

390	AGACCTCTCTCTTCTCTATGGAAGGGCCCACTTCATGAGGTGACACGACACAA	449
162	TGATTGTCTCTTCGCCATGAAAAGGCTGCCAACTTCATGATGACTACAGACAA	221
450	GTTACTCGGCATCTTTCAGAATGGGAATGTCTGTACAGATCAGCGCTGACCTCATTT	509
222	ATTCTCAAGATTTTCAAAAATGGAATGTCTTATTCATTAAGATTACATTAACT	281
510	GTCTCGCTGATGAGCCCAAGAACTTCCCATGAGACATCCAGCGTGCAGATGACCT	569
282	TTCTCTTCATCATGATCTCAGAAATTTTCCATGATGTACAAACATGTATTAATCACT	341
Db		
QY	570 TGAGAGCT 577	
Db	342 GGAAGACT 349	
RESULT 14		
LOCUS	AM280976	
DEFINITION	fj48d02.y1 zebrafish adult brain Danio rerio cDNA 5' similar to gb:x52009 GLYCINE RECEPTOR ALPHA-1 CHAIN PRECURSOR (HUMAN);, mRNA sequence.	
ACCESSION	AM280976	
VERSION	AM280976.1	
KEYWORDS	EST.	
SOURCE	zebrafish.	
ORGANISM	Danio rerio.	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes; Cyprinidae; Danio.	
AUTHORS	Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Stepec,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurr,R., Ritter,E., Kohn,S., Shih,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.	
TITLE	Washu zebrafish EST Project 1998	
JOURNAL	Unpublished (1998)	
COMMENT	Other_ESTs: fj48d02.x1 Contact: Stephen L. Johnson Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: zbratfish@watson.wustl.edu CDNA library Preparation: John Ngai. CDNA library Arrayed by: Matthew Clark. DNA Sequencing by: Washington University Genome Sequencing Center clone distribution: Genome Systems, St. Louis, Missouri (web address: www.genomesystems.com) (email contact: info@genomesystems.com) and Research Genetics, Huntsville, Alabama (web address: www.resgen.com) (email contact: info@resgen.com) and ResourceGenZentrumPrimatePatentbank, Berlin, Germany (web address: www.rzpd.de) Seq primer: T7	
FEATURES	Source	
	Location/Qualifiers	
	1..501	
	/organism="Danio rerio"	
	/db_xref="taxon:7955"	
	/clone_lib="zebrafish adult brain"	
	/sex="mixed male and female"	
	/tissue_type="brain"	
	/dev_stage="adult"	
	/lab_host="E. coli DH10B"	
	/note="Vector: pZIPLOX; site_1: NotI; site_2: SalI; Original library was constructed in lambdaZ1PLOX. Mass excision of the cDNA library was performed to yield pZIPLOX plasmids. Insert check was done in original library."	
BASE COUNT	114 a 137 c 123 g 127 t	
ORIGIN		

Query Match	10.3%;	Score 168.6;	DB 10;	Length 501;
Best Local Similarity	76.4%;	Pred. No. 2,4e-25;		
Matches 207;	Conservative 0;	Mismatches 64;	Indels 0;	Gaps 0;

QY

123

CATGTCCCTCTGTGATTTTCTAGACAACCTTAAGGGCCGAACATCTGGATATGATGCCAG

182

Db

231

CATGTCTCCCTCTGACTTCTCGATTAACAGATGGGACGAACCTCCGCTATGACGCTCG

290

QY

183

GATTCGGGCCCAATTTTAAAGGCCACCCCGGAACGTGACCTGGCAACATCTTCATCAACAG

242

Db

291

CATTGACACCAAACTTCAAGAGTCTCTCTGTGAAGCTCAGCTGATATATTATTCATCAACAG

350

QY

243

TTTCAGCTCCGCTACCAAGACACACATGAGTACTACGGGTGAATGCTTCTTTCGGCAACA

302

Db

351

CTTCGGTTCAAGTACAGAACTACTATAGGACTACAGGGTGAACATTTATCTCCGGCAAGAA

410

QY

303

GTGGAATGACCCACCGCTCTCTCTACCGAATATCTGATGACTCTCTGGACCTCGATCC

362

Db

411

GTGGAAGACGACCTCTGACTGGCTTCACAGAGATATCCGGAATCTGTTGGATCTGGACCC

470

QY

363

CTCCATGCTGACTCTATCTGAGGACCGACAG

393

Db

471

GTCCATGCTGACTCTCATCTGGAACCTGAC

501

RESULT\_15

BQ938794

975 bp

mRNA

linear

EST 21-AUG-2002

LOCUS

BQ938794

5', mRNA sequence.

5', mRNA sequence.

DEFINITION

BQ938794

AGNCOCURT\_8930535 NIH\_MGC\_94 Mus musculus cDNA clone IMAGE:6466531

ACCESSION

BQ938794

BQ938794.1 GI:22354272

VERSION

BQ938794

EST.

KEYWORDS

BQ938794

house mouse.

SOURCE

BQ938794

Mus musculus.

ORGANISM

BQ938794

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

BQ938794

1 (bases 1 to 975)

AUTHORS

BQ938794

NIH-MGC <http://imgc.nci.nih.gov/>.

TITLE

BQ938794

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

BQ938794

Unpublished (1999)

COMMENT

BQ938794

Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: The Cepko Laboratory  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Plate: LLAM13990 row: P column: 20  
High quality sequence stop: 638.

FEATURES

BQ938794

Location/Qualifiers

SOURCE

BQ938794

1..975

/organism="Mus musculus"

/db\_xref="taxon:10090"

/clone="IMAGE:6466531"

/clone\_1id="NIH\_MGC\_94"

/tissue\_type="retina"

/lab\_host="DH10B (phage-resistant)"

/note="Organ: eye; Vector: pCMV-SPORT6, Site.1: NotI; Site.2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 3.3 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH\_MGC library."

BASE COUNT

BQ938794

263 a. 234 c 221 g 257 t

ORIGIN

BQ938794

9.7%: Score 159.6; DB 14; Length 975;  
Best Local Similarity 54.3%: Pred. No. 1.6e-23;  
Matches 376; Conservative 0; Mismatches 274; Indels 42; Gaps 1

```
OY 343 GACTCTGTGACCTGATCCCTCCATGCTGACTCTATCTGAAACGACCTCTTCTT 402
  || ||| || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 64 GATGACACTGACAGTTGACCCACCATGATATAGTCTGTGTGAAACCTGACTTATCTT 123
OY 403 GCTAATGAGAAAGGGCCACTTCCATGAGGTGACACGACGACAACTACTGGCATT 462
  || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 124 GCAATGMAAAAGTCCAAATTTTCATGATGTGACCCAAAGAAATATCTGTGTTTATC 183
OY 463 TTCAGAAATGGGAATGCTGTACAGATCCAGGGTGACCCATTTTGTCCGTGATG 522
  || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 184 TTTGGGGATGAGAGCTCTTGTGAGCAAGGTGTCTTATTAACCTTCAATGCTCTA 243
OY 523 GACCTCAAGAACTTCCCATGAGACATCCAGAGTGCAGATGCAAGCTTGAGAGCTCATCC 582
  || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 244 GACTTAACCTGTGTTCCATGAGACACAGCTGCAAAATGCAACTGAGAGCT----- 298
OY 583 ATACTGTGACCCCTGTGCTCTGTCTGTCACTTTCAGTTGGCTTACCATCATGAAGACCTC 642
  || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 299 -----TTGGATATACAAACCGATATTTA 321
OY 643 GTGTTGAGTGGCTGGAAGATGCTGCTGTCCTCCAAAGTGCGTGAAGGGCTGACTGTGCC 702
  || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 322 AGATTCACTGTGGAGTCAAGAGATCTGTTCAGTTGMAAAATTTCTTACTCAATTT 381
OY 703 CAGTTTATCTTGCAGATGAGAGATCTAGGCTGTGTGTACCAAGCACTACACACAGG 762
  || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 382 GATATTTAAAGAGAGATATGGAATATGCAACTGTACAAATATCTATAAGGCACCTGT 441
OY 763 AAATTCACTGCTGAGAGTAAAGTTTCACTGCAAGCGCAGAGTGGCTACTATCTGATT 822
  || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 442 TACTACACTGTGTGAGGTATCTTCACTGAGAGACAGGTGGGTCTTACTATGATG 501
OY 823 CAGATGTACATCCAGCTACTCATCTCATCTGCTGCTGCTCTCTGATCAAC 882
  || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 502 GCGTATATGCAACCACTTACTGATGTGTCTCTCTGCTCTCTTCTGTGATCAAC 561
OY 883 ATGATGCTGCCCCCTGCTGTGTGGGCTGGGATACACACCGTGTCTCACCATGACAC 942
  || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 562 CCTGATGCTAGTGTGCGCAGAGTACCTCTGGGCACTTCTCGTGTCTCAGTTGGCTCA 621
OY 943 CAGAGCTGTGGCTCCCGGGCTCTTGGCTTAAGGTGTCTTACGTGAAGCAATGACATC 1002
  || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 622 GAGTGCACACCTCTGCGACGCGAGCTTCTAAAGTGTCTTAATGTGAAAGCGCTGATGTG 681
OY 1003 TGAATGCTGTGTGCTGCTCTTGTGTGCG 1034
  || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 682 TGGCTCATTTGCTGCTCTCTTGGGTTTGC 713
```

Search completed: June 30, 2003, 23:00:23  
Job time : 2198 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 25, 2003, 17:12:14 ; Search time 62.6734 Seconds  
(without alignments)  
916.354 Million cell updates/sec

Title: US-10-075-846-4  
Perfect score: 2252  
Sequence: 1 MTLVLPATLSFLLMLTLPQ.....POPPAPLRGERTTKRLVVD 431

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues  
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_101002:\*

1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*  
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*  
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2163	96.0	417	23	AAE21178 Human TRICH-22 pro
2	2157	95.8	417	21	AAE19336 Amino acid sequenc
3	871	38.7	496	23	ABE57052 Mouse ischemic co
4	754.5	33.5	517	22	AAE13035 R. sanguineus glut
5	742.5	33.0	450	22	AAE13037 R. sanguineus glut
6	742.5	33.0	450	22	AAE13038 R. sanguineus glut
7	737	32.7	444	23	AAE16395 Heliothis virescen
8	737	32.7	450	23	AAE16439 Heliothis virescen
9	736	32.7	537	23	AAE13312 Dermacentor variab
10	732	32.5	537	23	AAE13314 Dermacentor variab

11	731.5	32.5	455	20	AAW97860 Cat flea glutamate
12	731	32.5	537	23	AAE13313 Dermacentor variab
13	730.5	32.4	456	17	AAW05246 Drosophila glutama
14	729	32.4	606	22	ABE65121 Drosophila melanog
15	726	32.2	637	21	AAV51077 D. stimulans GABA I
16	725.5	32.2	462	22	AAE00863 Short form of S. a
17	725.5	32.2	462	22	AAE00864 Short form of S. a
18	724.5	32.2	459	20	GABA-gated chlorid
19	724.5	32.2	467	20	AAW81634 GABA-gated chlorid
20	724.5	32.2	481	20	AAW81635 GABA-gated chlorid
21	724.5	32.2	496	20	AAW81633 GABA-gated chlorid
22	720	32.0	488	19	AAW69285 GABA-gated chlorid
23	720	32.0	488	19	AAW69286 GABA-gated chlorid
24	720	32.0	606	23	AAE13315 Dermacentor melano
25	720	32.0	637	21	AAV51074 D. melanogaster po
26	720	32.0	637	21	AAV51075 D. melanogaster po
27	719.5	31.9	454	22	ABE63334 Drosophila melanog
28	719	31.9	397	22	AAE13034 Dermacentor variab
29	718	31.9	637	21	AAW9336 D. melanogaster GA
30	717	31.8	496	17	AAW89336 GABA receptor subu
31	717	31.8	496	20	AAW61611 Insect GABA-A recept
32	716	31.8	561	19	AAW97413 Lucilia cupripes GA
33	715	31.7	496	22	ABE65256 Drosophila melanog
34	712	31.6	474	14	AAE11188 GABA-A receptor de
35	712	31.6	474	15	AAE59866 Human GABA-A recept
36	704	31.3	637	14	AAE34035 Human polypeptide
37	698.5	31.0	478	22	AAW40884 Human polypeptide
38	692.5	30.8	474	22	AAW39098 Mouse ischemic co
39	692	30.7	473	23	ABE57078 Long form of S. am
40	690	30.6	513	22	AAE00861 Long form of S. am
41	690	30.6	513	22	AAE00862 Human gamma-amino
42	678.5	30.1	452	17	AAE97299 Mutant gamma-amino
43	674.5	30.0	487	23	AAU81289 Caenorhabditis ele
44	673	29.9	487	17	AAE88361 Caenorhabditis ele
45	670.5	29.8	510	17	AAE88360

## ALIGNMENTS

RESULT 1	AAE21178	AAE21178 standard; Protein: 417 AA.
ID	AAE21178:	
AC	AAE21178:	
XX		
DT	01-JUN-2002	(first entry)
XX		
DE	Human TRICH-22 protein.	
XX		
KW	Human; transporter and ion channel; TRICH-22; transport disorder; angina;	Cat flea glutamate
KW	amniotic lateral sclerosis; cystic fibrosis; neuromuscular disorder;	Dermacentor variab
KW	cardiac disorder; polymyositis; diabetes; neurological disorder; cancer;	Drosophila glutama
KW	depression; schizophrenia; anaemia; Wilson's disease; Cushing's disease;	Drosophila melanog
KW	cell proliferated disorder; infertility; arteriosclerosis; gene therapy;	D. stimulans GABA I
KW	Alzheimer's disease; Parkinson's disease; Huntington's disease; allergy;	Short form of S. a
KW	mysthenia gravis; multiple sclerosis; metabolic disorder; hypertension;	GABA-gated chlorid
KW	acquired immune deficiency syndrome; immunological disorder; scleroderma;	GABA-gated chlorid
KW	endocrine disorder; autoimmune thyroiditis; rheumatoid arthritis; goitre;	GABA-gated chlorid
KW	cardiac myopathy; amnesia; toxic myopathy; Addison's disease; infection;	GABA-gated chlorid
KW	epilepsy; mental disorder; myocarditis; Crohn's disease; Grave's disease;	Dermacentor melano
KW	muscle disorder; stroke; dementia; anxiety; AIDS; asthma; cirrhosis.	D. melanogaster po
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	Peptide	1..28
FT		/label- Signal_peptide
FT	Protein	29..417
FT		/note- "Mature human TRICH-22 protein"
FT	Domain	44..341
FT		/note- "Neurotransmitter-gated ion channel domain"
FT	Domain	172..186

FT Domain /note="Neurotransmitter-gated ion channel motif"  
FT 255..279  
FT /note="Transmembrane domain"  
FT 320..339  
FT Domain /note="Transmembrane domain"  
FT WO200212340-A2.  
XX 14-FEB-2002.  
XX 01-AUG-2001; 2001WO-US24217.  
XX 03-AUG-2000; 2000US-223269P.  
XX 10-AUG-2000; 2000US-224456P.  
XX 18-AUG-2000; 2000US-226410P.  
XX 25-AUG-2000; 2000US-228140P.  
XX 31-AUG-2000; 2000US-230067P.  
XX 08-SEP-2000; 2000US-231434P.  
XX (INCY-) INCYTE GENOMICS INC.  
XX  
XX Yue H, Thornton M, Ramkumar J, Tang YT, Azimzai Y, Baughn MR;  
PI Yang J, Yao MG, Lal P, Walla NK, Gandhi AR, Hafalia AJA;  
PI Nguyen DB, Patterson C, Elliott VS, Trilbouley CM, Lu DAM, Xu Y;  
PI Reddy R, Hernandez R, Borowsky ML, Lo TP, Lu Y, Policky JL;  
PI Greene BD, Sanjanwala MS, Raumann BE, Burford N, Ison CH, Lee EA;  
PI Ding L, Das D, Kallil Da, Khan FA, Sellhammer JJ;  
DR MPI: 2002-206330/26.  
DR N-PSDB: AAD33667.  
PT New human transporters and ion channels polypeptides and  
PT polynucleotides for diagnosing, preventing or treating transport,  
PT neurologicall, muscle, immunological and cell proliferative disorders  
XX  
PS Claim 66; Page 181-182; 230pp; English.

CC The invention relates to human transporter and ion channel polypeptides  
CC designated TRICH and nucleic acid molecules encoding such polypeptides.  
CC TRICH sequences are useful for diagnosis, treatment and prevention of  
CC transport, muscle, neurological, immunological and cell proliferative  
CC disorders. Transport disorders include akinesia, amphotrophic lateral  
CC sclerosis, ataxia telangiectasia, cystic fibrosis, Becker's muscular  
CC dystrophy, diabetes mellitus, diabetes insipidus, myasthenia gravis,  
CC myocardiitis, prostate cancer, cardiac disorders associated with  
CC transport e.g. polymyositis, bradyarrhythmia, dermatomyositis, angina,  
CC neurological disorders associated with transport e.g. amnesia, bipolar  
CC disorder, depression, Tourette's disorder, schizophrenia, other disorders  
CC associated with transport e.g. neurofibromatosis, sickle cell anaemia,  
CC Wilson's disease, cataracts, infertility, hypoglycaemia, hypoglycaemia,  
CC goitre, Cushing's disease, hypercholesterolaemia and cystinuria. Cell  
CC proliferated disorders include cancer, actinic keratosis, cirrhosis,  
CC arteriosclerosis, atherosclerosis, bursitis, hepatitis and psoriasis.  
CC Neurological disorders include Alzheimer's, Pick's and Parkinson's  
CC disease, amphotrophic lateral sclerosis, epilepsy, stroke, Huntington's  
CC disease, multiple sclerosis, dementia and other extrapyramidal disorder,  
CC motor neuron disorder, prion disease, metabolic disease of the nervous  
CC system and other developmental disorders of the central nervous system,  
CC neuromuscular disorders, metabolic, endocrine and toxic myopathies,  
CC periodic paralysis, mental disorders including mood, anxiety, and  
CC immunological disorders include acquired immune deficiency syndrome  
CC (AIDS), adult respiratory distress syndrome, Addison's disease,  
CC allergies, asthma, atherosclerosis, osteoporosis, autoimmune haemolytic  
CC anaemia, autoimmune thyroiditis, Crohn's disease, atopic dermatitis,  
CC Grave's disease, glomerulonephritis, rheumatoid arthritis, scleroderma,  
CC systemic lupus erythematosus, systemic sclerosis, ulcerative colitis,  
CC haemodialysis, uveitis; viral, bacterial, fungal, parasitic, protozoal,  
CC helminthic infections and trauma; and muscle disorders include cardiac  
CC myopathy, myocarditis, polymyositis, arrhythmias and hypertension. The  
CC TRICH polynucleotides are used in gene therapy. The present sequence  
CC is human TRICH-22 protein.  
XX  
XX Sequence 417 AA:

Query Match 96.0%; Score 2163; DB 23; Length 417;  
Best Local Similarity 96.5%; Pred. No. 1,4e-216;  
Matches 416; Conservative 0; Mismatches 1; Indels 14; Gaps 1;  
QY 1 MTTLVPAATLSELLMTLTGQVLLRVALAKEEVKSGTSGSPSPDFLDKLMGRTSGYDA 60  
DB 1 MTTLVPAATLSELLMTLTGQVLLRVALAKEEVKSGTSGSPSPDFLDKLMGRTSGYDA 60  
QY 61 RIRPFKGPVNVNCFINFSSTVTKTMDYRVNVPFRQWNPRLSYREYPPDSLDLD 120  
DB 61 RIRPFKGPVNVNCFINFSSTVTKTMDYRVNVPFRQWNPRLSYREYPPDSLDLD 120  
QY 121 PSMDSIRKPDLEFANEKANFHEVTDNKLIRFKNGNVLYSIRLTILISCLMDLKNFP 180  
DB 121 PSMDSIRKPDLEFANEKANFHEVTDNKLIRFKNGNVLYSIRLTILISCLMDLKNFP 180  
QY 181 MDIOTCTMOLESSSILSPPLSLSLVGYTKMDLFEWLEDAPAVQVAEGTLTLPQFILRD 240  
DB 181 MDIOTCTMOLESSSILSPPLSLSLVGYTKMDLFEWLEDAPAVQVAEGTLTLPQFILRD 240  
QY 241 EKDIACCCKHNTGKFTGIEYKPHLEROMGYLLQMTIPSLIYILSVSWFMWMDAPA 300  
DB 227 EKDIACCCKHNTGKFTGIEYKPHLEROMGYLLQMTIPSLIYILSVSWFMWMDAPA 286  
QY 301 RVGLGITTVMITTOSSSGSRASLPKVSIVKADIDIMAVACLFVPAALLEVAIAINFSROH 360  
DB 287 RVGLGITTVMITTOSSSGSRASLPKVSIVKADIDIMAVACLFVPAALLEVAIAINFSROH 346  
QY 361 KEFIRRRRRORRLEEDIIODESRFPGYGLGHCLARDGCPMEGSGTISYPOPPAPLLR 420  
DB 347 KEFIRRRRRORRLEEDIIODESRFPGYGLGHCLARDGCPMEGSGTISYPOPPAPLLR 406  
QY 421 EGETTRKLYVD 431  
DB 407 EGETTRKLYVD 417

RESULT 2  
AAB19336  
ID AAB19336 standard; Protein; 417 AA.  
XX  
XX AAB19336:  
AC 19-FEB-2001 (first entry)  
XX  
XX 19-FEB-2001 (first entry)  
DE Amino acid sequence of a human ataxia protein.  
XX  
XX Human; ataxia; gene therapy.  
XX  
XX Homo sapiens.  
OS WO200058461-A1.  
XX  
XX 05-OCT-2000.  
PD 23-MAR-2000; 2000WO-EP02600.  
XX  
XX 26-MAR-1999; 99EP-0106343.  
PR (RAPP/) RAPPOLD-HOERBRAND G.  
XX  
XX RAPPOLD-HOERBRAND G;  
PI MPI: 2000-656166/63.  
XX N-PSDB: AAC61678.  
DR Novel nucleic acid sequence encoding human ataxia protein for screening  
PT compounds useful for treating disorders relating to mutations in ataxia  
PT gene  
XX  
XX Claim 10; Page 20-21; 47pp; English.



CC The present sequence represents a human ataxia protein. The ataxia  
CC protein and polynucleotides are useful for diagnosing and treating  
CC disorders related to ataxia. Ataxia gene sequences are useful in  
CC gene therapy, and as diagnostic tools or reagents for identifying and  
CC characterizing genetic defect involved in the disorders and diseases  
CC related to ataxia.

XX Sequence 417 AA;

Query Match 95.8%; Score 2157; DB 21; Length 417;  
Best Local Similarity 96.1%; Pred. No. 3, 7e-216;  
Matches 414; Conservative 1; Mismatches 2; Indels 14; Gaps 1;

QY 1 MTLVATISFLMLTLPQVLLRVALAKEEVSGTSGSPSPDLKMGRTSGYDA 60  
DB 1 MTLVATISFLMLTLPQVLLRVALAKEEVSGTSGSPSPDLKMGRTSGYDA 60  
QY 61 RIRPNKGPVNVNTCNIFINSFSSVTKTMDYVNVFLRQNDPRLSTREYDDSLDD 120  
DB 61 RIRPNKGPVNVNTCNIFINSFSSVTKTMDYVNVFLRQNDPRLSTREYDDSLDD 120  
QY 121 PSMUDSIMKPDLEFANEKANHEVTDNKLRIFKNGANVLSIRLTLLISCLMDLKNP 180  
DB 121 PSMUDSIMKPDLEFANEKANHEVTDNKLRIFKNGANVLSIRLTLLISCLMDLKNP 180  
QY 181 MDIOTCTMOLESSSILCSPLPSLSVGYTMKDLVEFMELEDAVAVOAEGLTLPOLLRD 240  
DB 181 MDIOTCTMOLESSSILCSPLPSLSVGYTMKDLVEFMELEDAVAVOAEGLTLPOLLRD 240  
QY 241 EKDLGCTTKHYNTGKFTCEVFKFLEROMGYLIQWYIPSLIIVLSWSEWIMDAAPA 300  
DB 241 EKDLGCTTKHYNTGKFTCEVFKFLEROMGYLIQWYIPSLIIVLSWSEWIMDAAPA 300  
QY 227 EKDLGCTTKHYNTGKFTCEVFKFLEROMGYLIQWYIPSLIIVLSWSEWIMDAAPA 286  
DB 227 EKDLGCTTKHYNTGKFTCEVFKFLEROMGYLIQWYIPSLIIVLSWSEWIMDAAPA 286  
QY 301 RVGLGTTVLTMTQSSGRASLPKVSYYKAIIDIMAVCLLFYFAALLEYAAINPVSROH 360  
DB 301 RVGLGTTVLTMTQSSGRASLPKVSYYKAIIDIMAVCLLFYFAALLEYAAINPVSROH 360  
QY 287 RVGLGTTVLTMTQSSGRASLPKVSYYKAIIDIMAVCLLFYFAALLEYAAINPVSROH 346  
DB 287 RVGLGTTVLTMTQSSGRASLPKVSYYKAIIDIMAVCLLFYFAALLEYAAINPVSROH 346  
QY 361 KEFIRLRORRORLEEDIIQESRFYFGYGLCHCLQARDGPMESGSIYPOPPAPLLR 420  
DB 361 KEFIRLRORRORLEEDIIQESRFYFGYGLCHCLQARDGPMESGSIYPOPPAPLLR 420  
QY 347 KEFIRLRORRORLEEDIIQESRFYFGYGLCHCLQARDGPMESGSIYPOPPAPLLR 406  
DB 347 KEFIRLRORRORLEEDIIQESRFYFGYGLCHCLQARDGPMESGSIYPOPPAPLLR 406  
QY 421 EGETTRKLYVD 431  
DB 421 EGETTRKLYVD 431  
QY 407 EGETTRKLYVD 417  
DB 407 EGETTRKLYVD 417

RESULT 3

ID ABB57052 standard; Protein: 496 AA.

XX ABB57052;

DT 07-MAR-2002 (first entry)

DE Mouse ischaemic condition related protein sequence SEQ ID NO:86.

XX Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;  
KM vasospastic ischaemia; ischaemic condition; ischaemic disease.

XX Mus musculus.

XX MO200188188-A2.

XX 22-NOV-2001.

PE 18-MAY-2001; 2001MO-JP04192.

PR 18-MAY-2000; 2000JP-0145977.

PA (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.

XX Ishikawa K, Asei S, Takahashi Y, Nagata T, Ishii Y;

DR WPI; 2002-034733/04.

DR N-PSDB; AB199254.

PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring

PT expression levels of particular genes defined in the specification or

PS Claim 2; Page 269-271; 2690pp; English.

CC The present invention describes a method for examining ischaemic  
CC conditions, comprising measuring the expression levels of particular  
CC genes (I) in a test sample or determining the expression profile of a  
CC gene group in the sample comprising genes selected from (I). The method  
CC is useful for examining the ischaemic condition (e.g. compressive  
CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring  
CC expression levels of particular genes (AB199202 to AB199912, encoding  
CC the protein sequences in ABB57020 to ABB57374) or by determining the  
CC expression profile of a gene group comprising these genes. The  
CC expression levels or expression profiles produced by these genes are  
CC used as an indicator when screening for ischaemic condition-improving  
CC drugs or therapeutics for ischaemic diseases. AB199913 and AB199914  
CC represent PCR primers for a mouse ischaemic condition related sequence,  
CC which are used in the exemplification of the present invention.

XX Sequence 496 AA;

Query Match 38.7%; Score 871; DB 23; Length 496;  
Best Local Similarity 48.5%; Pred. No. 1e-81;  
Matches 188; Conservative 61; Mismatches 87; Indels 52; Gaps 11;

QY 9 LSFLLMTLPQVLLRVALAKEEVSGTSGSQSPDLARVPNSTNLNR 62  
DB 9 LSFLLMTLPQVLLRVALAKEEVSGTSGSQSPDLARVPNSTNLNR 62  
QY 7 ISFILLMSL-----LFEDACKKESKKKQYLCPSQSPDLARVPNSTNLNR 107  
DB 7 ISFILLMSL-----LFEDACKKESKKKQYLCPSQSPDLARVPNSTNLNR 107  
QY 51 LMGRTGYDARIRPNKGPVNVNTCNIFINSFSSVTKTMDYVNVFLRQNDPRL--- 119  
DB 51 LMGRTGYDARIRPNKGPVNVNTCNIFINSFSSVTKTMDYVNVFLRQNDPRL--- 119  
QY 63 LL---VSYDRIIRPNKGPVNVNTCNIFINSFSSVTKTMDYVNVFLRQNDPRL--- 166  
DB 63 LL---VSYDRIIRPNKGPVNVNTCNIFINSFSSVTKTMDYVNVFLRQNDPRL--- 166  
QY 108 -SYREYDDSLDDPSMUDSIMKPDLEFANEKANHEVTDNKLRIFKNGANVLSIRL 177  
DB 108 -SYREYDDSLDDPSMUDSIMKPDLEFANEKANHEVTDNKLRIFKNGANVLSIRL 177  
QY 120 SDFR--GSDALVDPMTWKLMKPDLEFANEKANHEVTDNKLRIFKNGANVLSIRL 222  
DB 120 SDFR--GSDALVDPMTWKLMKPDLEFANEKANHEVTDNKLRIFKNGANVLSIRL 222  
QY 167 TLLISCLMDLKNPMDIOTCTMOLESSSILCSPLPSLSVGYTMKDLVEFMELEDAVAVO 226  
DB 167 TLLISCLMDLKNPMDIOTCTMOLESSSILCSPLPSLSVGYTMKDLVEFMELEDAVAVO 226  
QY 178 SITLSCPLDLPFMDQORCKMQLE-----SFGYTDDLRIMQSGD-VQ 222  
DB 178 SITLSCPLDLPFMDQORCKMQLE-----SFGYTDDLRIMQSGD-VQ 222  
QY 227 VAEGLTLPQVLLRVALAKEEVSGTSGSQSPDLARVPNSTNLNR 284  
DB 227 VAEGLTLPQVLLRVALAKEEVSGTSGSQSPDLARVPNSTNLNR 284  
QY 223 L-EKIALPOFDIKEDIEYGNCTKYYKGTGYTCVEVIFTLRROVGYMMGVAPYAPLLIY 281  
DB 223 L-EKIALPOFDIKEDIEYGNCTKYYKGTGYTCVEVIFTLRROVGYMMGVAPYAPLLIY 281  
QY 285 ILSWSEWIMDAAPAVGIGITVLTMTQSSGRASLPKVSYYKAIIDIMAVCLLFYF 344  
DB 285 ILSWSEWIMDAAPAVGIGITVLTMTQSSGRASLPKVSYYKAIIDIMAVCLLFYF 344  
QY 282 VLSWSEWIMDAAPAVGIGITVLTMTQSSGRASLPKVSYYKAIIDIMAVCLLFYF 341  
DB 282 VLSWSEWIMDAAPAVGIGITVLTMTQSSGRASLPKVSYYKAIIDIMAVCLLFYF 341  
QY 345 AALLEYAAINPVSROHKEFIRLRORR 372  
DB 345 AALLEYAAINPVSROHKEFIRLRORR 372  
QY 342 ASLVEYAVVQVMLNPK---RVEAEKRR 366  
DB 342 ASLVEYAVVQVMLNPK---RVEAEKRR 366

RESULT 4

ID AAE13039 standard; Protein: 517 AA.

XX AAE13039;

DT 28-JAN-2002 (first entry)

DE R. sanguineus glutamate-gated chloride channel 1 protein, T32.

XX Brown dog tick; glutamate-gated chloride channel; GluCl1;

KW GluCl2; crop protection; insecticide; nematocide; acaricide;  
KM T32 protein.

```

XX  Rhipicephalus sanguineus.
OS
XX
XX  WO200174838-A1.
XX
XX  11-OCT-2001.
XX
XX  28-MAR-2001; 2001WO-US09905.
XX
XX  31-MAR-2000; 2000US-193934P.
XX
XX  (MERI ) MERCK & CO INC.
XX
XX  Warmke JW, Yang Y, Cully DF, Hamelin MJ;
XX
XX  WPI: 2001-662963/76.
XX  N-PSDB: AAD21397.
XX
XX  Novel L-glutamate-gated chloride channel proteins from Rhipicephalus
XX  sanguineus for identifying compounds which modulate the channel
XX  proteins, which are useful as insecticides, anthelmintics and
XX  acaricides.
XX
XX  Claim 32; Fig 6; 89pp: English.
XX
XX  The invention relates to Rhipicephalus sanguineus (brown dog tick)
XX  L-glutamate-gated chloride channel proteins (Gluc11 and Gluc12)
XX  and nucleic acid molecules encoding such proteins. Gluc1 channel
XX  proteins are useful for identifying modulators. The compounds
XX  identified as modulators are useful for insecticidal, mitacidal
XX  and/or nematocidal treatment for use in animal and human health
XX  for and selecting compounds active against parasitic invertebrate
XX  species relevant to animal and human health, including worms,
XX  fleas, ticks, mites and lice. Heterologous cell lines expressing
XX  functional Gluc11 and Gluc12 channel functional forms are useful
XX  for establishing functional or binding assays to identify novel
XX  Gluc11 protein, T12.
XX
XX  Sequence 517 AA:
SQ
Query Match 33.5%; Score 754.5; DB 22; Length 517;
Best Local Similarity 39.9%; Pred. No. 1.5e-69;
Matches 180; Conservative 69; Mismatches 133; Indels 69; Gaps 16;
OY 5 VP-ATLSFLL-----W-TLP-----GGVLRVLA-----KEEYKSGTGS 39
    || || || || || || || || || || || || || || || || || || || ||
DB 10 VFLVALAFPLLILSCPSMAETLPTPPRGQGVVAAMLLGKOQSSRIODKEG-RAN 68
    || || || || || || || || || || || || || || || || || || || ||
OY 40 QPMSPDELDKLMGRTSGYDARIRP-----NEKGPVNYTCNIFINSFSSVTKTMDRYN 95
    || || || || || || || || || || || || || || || || || || || ||
DB 69 FRAIEKRLDSIIQG-GRYDCRIRPMGINNTDGPAL-VRVNIFFVRSIGRIDVTMEYTI 126
    || || || || || || || || || || || || || || || || || || || ||
OY 96 VFLRQWMDPRLSYREYDDSDLDPSMLDSIWKDKLFPANEGANFHEVTTDNKLRIF 155
    || || || || || || || || || || || || || || || || || || || ||
DB 127 MTFRQWMDPRLSYREYDDSDLDPSMLDSIWKDKLFPANEGANFHEVTTDNKLRIF 186
    || || || || || || || || || || || || || || || || || || || ||
OY 156 KNGNVLYSIRLLILSLCMLDKNFPMDIOTQMOLESSILCSPLSLSVGYTMKIDV 215
    || || || || || || || || || || || || || || || || || || || ||
DB 187 PNGDVLFSIRISLVSCPMNLKFYPLDKOCSI-----VWASGYTTEDLV 232
    || || || || || || || || || || || || || || || || || || || ||
OY 216 FEMLEDAVAVOAEGLTLPQFTL-RDEKDLGCTFKNYTKGKTCLEVFHLERQGYLI 274
    || || || || || || || || || || || || || || || || || || || ||
DB 233 FLMKRGDP-VQYTKMLHLPFRFTLRFQTDY--CTSRVTMGESVCLRVDLVFRRESYLI 289
    || || || || || || || || || || || || || || || || || || || ||
OY 275 QMTITSLIIVISVSWINMDAARVGLGITYLVMTQSSGSRASLPKSYKALDI 334
    || || || || || || || || || || || || || || || || || || || ||
DB 290 QIYICCMIVYSWVSMFLDPTSIAPARVSLGVTTLTMTATQISGINASLPVSYKADIV 349
    || || || || || || || || || || || || || || || || || || || ||
OY 335 WAAVCLLEFVALLEYAALNFEVSRQHKFEIRLRQRQRO-----RLEEDIQESFFYRG 389
    || || || || || || || || || || || || || || || || || || || ||
DB 350 WTGVCLTFEFGALFEFALVNTASRSDSRQNMOKOKORKEWLEPLDSDHEDGATTPAM 409

```

```

OY 390 YGLGHCLQARDCGPMESGIVS---POPPAP 417
DB 410 VSSG-----EPAGLMAKRTWPPPLP 429
XX
XX  RESULT 5
XX  ID AE13037 standard; Protein; 450 AA.
XX  AC AE13037;
XX  DT 28-JAN-2002 (first entry)
XX
XX  R. sanguineus glutamate-gated chloride channel 1 protein, T12.
XX
XX  Brown dog tick; glutamate-gated chloride channel; Gluc11;
XX  Gluc12; crop protection; insecticide; nematocide; acaricide;
XX  T12 protein.
XX
XX  Rhipicephalus sanguineus.
XX
XX  WO200174838-A1.
XX
XX  11-OCT-2001.
XX
XX  28-MAR-2001; 2001WO-US09905.
XX
XX  31-MAR-2000; 2000US-193934P.
XX
XX  (MERI ) MERCK & CO INC.
XX
XX  Warmke JW, Yang Y, Cully DF, Hamelin MJ;
XX
XX  WPI: 2001-662963/76.
XX  N-PSDB: AAD21395.
XX
XX  Novel L-glutamate-gated chloride channel proteins from Rhipicephalus
XX  sanguineus for identifying compounds which modulate the channel
XX  proteins, which are useful as insecticides, anthelmintics and
XX  acaricides.
XX
XX  Claim 26; Fig 2; 89pp: English.
XX
XX  The invention relates to Rhipicephalus sanguineus (brown dog tick)
XX  L-glutamate-gated chloride channel proteins (Gluc11 and Gluc12)
XX  and nucleic acid molecules encoding such proteins. Gluc1 channel
XX  proteins are useful for identifying modulators. The compounds
XX  identified as modulators are useful for insecticidal, mitacidal
XX  and/or nematocidal treatment for use in animal and human health
XX  for and selecting compounds active against parasitic invertebrate
XX  species relevant to animal and human health, including worms,
XX  fleas, ticks, mites and lice. Heterologous cell lines expressing
XX  functional Gluc11 and Gluc12 channel functional forms are useful
XX  for establishing functional or binding assays to identify novel
XX  Gluc11 protein, T12.
XX
XX  Sequence 450 AA:
SQ
Query Match 33.0%; Score 742.5; DB 22; Length 450;
Best Local Similarity 44.0%; Pred. No. 2.2e-68;
Matches 157; Conservative 61; Mismatches 110; Indels 29; Gaps 8;
OY 48 LDKLMGRTSGYDARIRP-----NEKGPVNYTCNIFINSFSSVTKTMDRYNVLROQWN 103
    || || || || || || || || || || || || || || || || || || || ||
DB 42 LDSIIQG-GRYDCRIRPMGINNTDGPAL-VRVNIFFVRSIGRIDVTMEYTIQMTFREQWR 99
    || || || || || || || || || || || || || || || || || || || ||
OY 104 DPRLSYREYDDSDLDPSMLDSIWKDKLFPANEGANFHEVTTDNKLRIFKNGNVLXS 163
    || || || || || || || || || || || || || || || || || || || ||
DB 100 DERLOYDDLGQGVRYLTLPEDKLMKPDLPFSNEKEGFHNIIMPVLLRIHPNGDVLFS 159

```

Oy		164	IRLTLLITSLMDKPNPMDIOICTQOLSSSILSPSLSTLGYMKDLPFWEIDAP	223
Dd		160	IRISLVLSCEPMKNKFPLDKQICSI-----VWVSIGTTEDLVFLMKEGDP	205
Oy		224	AVOVAEGTLTPOFIL-RDEKDGLCCCTKHYNTEGFECIEVKFHLRONGYLLIOMYTIPSL	282
Dd		206	-VVTKNLHLPRETLERFGTDY--CTSRTNNGEYSCLRLVDLVFKREFRSYYLIQIYPCSM	262
Oy		283	IYILSVSWFWINMDAAPARVAGGITVTYLTMTQSSGSASLPKSYKYKAIDIMAVCLLF	342
Dd		263	LTVLSVWSFWLDDPTSLPARVSLGVVTLTLMATQISGINASLPVPSYRKADIVDWGCGLTF	322
Oy		343	VFAALLLEYAINEFSROHEFTFLRRORRQ-----RLEEDIIOESRFYRGYGLGH	394
Dd		323	VFEALLEGPLVLNVASHSDSRQNMOKOKORKWELPEPLDSOHLDEGATTAMRLPVH	379
	RESULT 6			
	AAEI3038	ID	AAEI3038 standard; Protein; 450 AA.	
	XX	AC	AAEI3038;	
DT	XX	AEI3038:		
DT	XX	28-JAN-2002	(first entry)	
De	XX	R. sanguineus	glutamate-gated chloride channel 1 protein, T82.	
KW	XX	Brown dog tick;	glutamate-gated chloride channel; GluC11;	
KW	XX	Gluc12:	crop protection; insecticide; nematocide; acaricide;	
KW	XX	T82 protein.		
OS	XX	Rhipicephalus	sanguineus.	
PX	XX	MO200174838-A1.		
PD	XX	11-OCT-2001.		
FZ	XX	28-MAR-2001;	2001WO-US09905.	
PR	XX	31-MAR-2000;	2000US-193934P.	
PA	XX	(MERI )	MERCK & CO INC.	
PI	XX	Warmke JW,	Yang Y, Cully DF, Hamelin MJ;	
DR	XX	WPI; 2001-662963/76.		
DR	XX	N-PSDB; AAD21396.		
PT	XX	Novel L-glutamate-gated	chloride channel proteins from Rhipicephalus	
PT	XX	sanguineus for identifying	compounds which modulate the channel	
PT	XX	acaricides -	which are useful as insecticides, anthelmintics and	
PS	XX	Claim 29; Fig 4;	89pp: English.	
CC	XX	The invention relates to	Rhipicephalus sanguineus (brown dog tick)	
CC	XX	L-glutamate-gated chloride	channel proteins (GluC11 and GluC12)	
CC	XX	and nucleic acid molecules	encoding such proteins. GluC1 channel	
CC	XX	proteins are useful for	identifying modulators. The compounds	
CC	XX	identified as modulators are	useful for insecticidal, mitacidal	
CC	XX	and/or nematocidal treatment	for use in animal and human health	
CC	XX	and/or crop protection. The	compounds are also useful in screening	
CC	XX	for and selecting compounds	active against parasitic invertebrate	
CC	XX	species relevant to animal	and human health, including worms,	
CC	XX	fleas, ticks, mites and lice.	Heterologous cell lines expressing	
CC	XX	functional Gluc11 and Gluc12	channel functional forms are useful	
CC	XX	for establishing functional or	binding assays to identify novel	
CC	XX	Gluc1 channel modulators. The	present sequence is R. sanguineus	
CC	XX	GluC11 protein, T82.		
Sequence	450 AA:			

33.0%; Score 742.5; DB 22; Length 450;

Best Local Similarity	44.0%	Pred. No. 2.2e-68	Matches 157	Conservative 61	Mismatches 110	Indels 29	Gaps 8
Qy	48 LDKLMGRTSGYDARI	---NFKGPPVNTCNIFINSFSSVTKTMDYRVNVFLRQON	103				
Db	42 LDSIQG-GRYDCRIRPG	INNTDGPAL-VRVNIFVRSIGRIDVDTMEYTVQMTFREQMR	99				
Qy	104 DPLRSTREYPPDSD	LDPSMDSIMKPRLPFRANEKANFHEVYTDNKLRLFKNGNVL	163				
Db	100 DERLOQYDLDGQV	RLTLTEPDKLMKPDLPFSNEKEGHNNIMPNNVLLRIHPNCDVLS	159				
Qy	164 IRLTLISCLMDLKN	FPMDIOTCTMQMLESSTLCSPLPSLSVGYTMKDLVFEWLEDP	223				
Db	160 IRISLVLSGCPMNL	KRYPDLKOICST-----VMSYGYTTEDLVFLMKEGDP	205				
Qy	224 AVQVAGEITLPOFIL	-RDEKDLGCGCTKHYNQKFTCEIEVKFHLEROMGYLLIOMYIPSL	282				
Db	206 -VQYTKNLHLPR	FTLERFQTOY--CTSTNTNGEYSCLVADLVFKREFSYLQIYIPCCM	262				
Qy	283 IVILSWNSFWINMD	APARVGLGITVYLMTTQSSGSNASLPKYSVKAIDIMAVCLLF	342				
Db	263 LVIVSWNSFWMDP	TSIPASVLSGYTTLTLMATQISINASLPVSYTKAIDWMTGCLTF	322				
Qy	343 VFALLLEVAALNF	PSRKHKEFRLRRRRQ----RLIEDIORSRYFFRCYGLGH	394				
Db	323 VEGALLLEVALNV	ASRSDSRNNOKOKORKEWLEPPIDSDHLEDGATTFAMRPLVH	379				
RESULT 7							
AAEL6395	AAEL6395 standard; Protein: 444 AA.						
AAEL6395							
AC	AAEL6395:						
DT	09-APR-2002 (first entry)						
xx	Heliothis virescens glutamate-gated chloride channel.						
xx	Lepidopteran glutamate-gated chloride channel; insecticide.						
xx	Heliothis virescens.						
xx	Key	Location/Qualifiers					
FT	Misc-difference 12..15	/note= "Encoded by TTGGTCTGTGA"					
FT	Misc-difference 207	/note= "Encoded by AAN"					
FT	Misc-difference 280	/note= "Encoded by TCA"					
FT	Misc-difference 327	/note= "Encoded by TTC"					
FT	Misc-difference 328	/note= "Encoded by GCG"					
FT	Misc-difference 329	/note= "Encoded by CTC"					
xx	US6329174-B1.						
xx	11-DEC-2001.						
xx	13-JUN-2000; 2000US-0592891.						
xx	13-JUN-2000; 2000US-0592891.						
xx	(AVET ) AVENTIS CROPS	SCIENCE SA.					
xx	Wang XM, Sarda XG,	Tomalski MD, Wingate VPM;					
xx	WPI; 2002-121133/16.						
xx	N-PSDB; AAD26939.						
xx	New nucleic acid encoding lepidopteran chloride channel, useful for screening agents for insecticidal activity -						

PS	XX	Claim 1; Column 14; 18pp; English.
CC	XX	The invention relates to nucleic acid encoding lepidopteran
CC	XX	glutamate-gated chloride channel. Glutamate-gated chloride channels
CC	XX	are a family of ligand-gated chloride channels unique to invertebrates.
CC	XX	The DNA of the invention is used for recombinant production of
CC	XX	lepidopteran glutamate-gated chloride channel and this is used,
CC	XX	optionally in the form of membrane preparations or recombinant cells,
CC	XX	in specific binding or functional assays for identifying potential
CC	XX	insecticides. The present sequence is <i>Heliothis virescens</i>
CC	XX	glutamate-gated chloride channel.
CC	XX	Note: This sequence SEQ.ID.NO.14 is stated to be similar to the
CC	XX	sequence shown in the sequence listing. However this sequence lacks
CC	XX	N-terminal six residues of the sequence shown in sequence listing.
SO	XX	Sequence 444 AA;
SO	XX	Query Match 32.7%; Score 737; DB 23; Length 444;
SO	XX	Best Local Similarity 45.7%; Pred.No. 8.2e-68;
SO	XX	Matches 158; Conservative 53; Mismatches 11; Indels 24; Gaps 7;
OY	DB	31 EVKSGTGSGQMSQMSDFLDKLMGRSSGDAIRP---NFKGPVAVTGNITFNSSSVTK 87
OY	DB	20 ECGMGKINFEKEKEKQIIDQILG-PGRYDARIRPSGINGTGPVAVSVANIFVRSISKIDD 78
OY	DB	88 TTMDYRVAVFPRQOMNDPRLSYREYRDPDSLDDPSMDSIMKPDLPFANEKGANHEVTT 147
OY	DB	79 VTMEYSVGLTFREQMDLDERLKFNNLGGRLKLTITLEARNVAMPDLFFENKEGHNITIM 138
OY	DB	148 DNKLIRFKNNVLYSTRITLTLSCMDLKNFMDIOTCTQMOLESSITLSPSLISV 207
OY	DB	139 PNVIYRIFFPNNNVYSIRISLTLSCPNMLKLYPLDKQYCSLRM-----ASY 184
OY	DB	208 GYTKMDLFEWLEDPAPAVQVAGEGLTLPOFILRDEKDL-GCCTKHNTGKFTICIEYKPHLE 266
OY	DB	185 GMTDDLDFLMEKEDP-VQYVKNLHLPRFTL--EKELFDYCNSTKNTGEYGLKYDLFLK 241
OY	DB	267 ROMGYLIQMTIPSLLIYILSVSWFWMIMDAPARVAGITTVTLTMTQSSGSRASLPKV 326
OY	DB	242 REFESYLIQIYIPCCMEIYVSWVFEMLDQGVAPRVLLGVTTLTLMATQSSGINASLPV 301
OY	DB	327 SYVAIDIMMAVCLLEFPAFLLEVAALINFAVR--QHKEFITLRRO 370
OY	DB	302 SYTAIDVMTGVCLETFVFGALLSEFVYASRSDMHRENKKARRE 347
RESULT 8		
AAEL16439		
ID	AAEL16439	standard; Protein; 450 AA.
XX	XX	AAEL16439;
AC	XX	
DT	DT	09-Apr-2002 (first entry)
DE	XX	<i>Heliothis virescens</i> glutamate-gated chloride channel, alternative form.
XX	XX	Lepidopteran glutamate-gated chloride channel; insecticide.
OS	OS	<i>Heliothis virescens</i> .
XX	XX	Key Location/Qualifiers
FT	FT	Misc-difference 1..7
FT	FT	Misc-difference 18..21
FT	FT	Misc-difference 213
FT	FT	Misc-difference 286
FT	FT	Misc-difference 333
FT	FT	Misc-difference 334

FT	Misc-difference	335	/note= "Encoded by GCG"
FT			"Encoded by CTC"
XX			
PX	US6329174-B1.		
PN			
PD			
PP	11-DEC-2001.		
XX			
PF	13-JUN-2000; 2000US-0592891.		
XX			
PR	13-JUN-2000; 2000US-0592891.		
XX			
PA	(AVET ) AVENTIS CROPS SCIENCE SA.		
PI	Wang XM, Sarda XG, Tomalski MD, Wingate VPM;		
XX			
DR	WPI; 2002-121133/16.		
DR	N-PSDB; AAD26939.		
XX			
PT	New nucleic acid encoding lepidopteran chloride channel, useful for		
PT	screening agents for insecticidal activity -		
PS			
XX	Claim 1; Column 27-30; 18pp; English.		
CC	The invention relates to nucleic acid encoding lepidopteran		
CC	glutamate-gated chloride channel. Glutamate-gated chloride channels		
CC	are a family of ligand-gated chloride channels unique to invertebrates.		
CC	The DNA of the invention is used for recombinant production of		
CC	lepidopteran glutamate-gated chloride channel and this is used,		
CC	optionally in the form of membrane preparations or recombinant cells,		
CC	in specific-binding or functional assays for identifying potential		
CC	insecticides. The present sequence is Heliothis virescens		
CC	glutamate-gated chloride channel, alternative form.		
CC	Note: This sequence SEQ.ID.NO.14 is stated to be similar to the		
CC	sequence shown in column 14 of the specification. However this		
CC	sequence contains six additional residues at the N-terminal end.		
XX			
SQ	Sequence	450 AA;	
	Query Match	32.7%; Score 737; DB 23; Length 450;	
	Best Local Similarity	45.7%; Pred. No. 8,3e-68;	
	Matches	158; Conservative 53; Mismatches 111; Indels 24; Gaps 7;	
OY	31 EVKSGTGSGPMSPSDLDKMGRTSGTDARIP--NFKGPVVNTCNIFNSFSVTK	87	
Dd	-26 ECMNGGKINFREKEKQILDQILG-PGRDARIKPSGINGTGPVAVSVNIFFRSISKIDD	84	
OY	88 TTMDYRVNVFLROOMNDPLSYREVPDDSLDLPEDMLDSIKPKDLFEFANEGANHEVT	147	
Dd	85 VTMEYSVOQLTRBOMLDERLKFNNLGGKLKLTITLEANRWMPDLFEFSNEKEGHNNIM	144	
OY	148 DNKLRLRFKNGNVLVSIRLLTLSCLDMLDNKFNPDIOQTQNOLESSILCSPLSPSLSV	207	
Dd	145 PNVVIRIDLFEPNNGVIVSIRISLTLSOPMKLYLPDLKQCSDLRM-----ASY	190	
OY	208 GTYKKDLVFEELEDPAPAOVAEGLTPOLFLRDEKD-LGCCSKHNTEGKFCIEYKFHLE	266	
Dd	191 GMTVDLDLVFLMKESDP-VQVVKNHHLRPFTL-KEFLLDYCNKSNTGTGYSLCKAYDLFK	247	
OY	267 ROMCYUJLOMYIPLSLIAVLSWFSFWIMDAAPARAGGITVLTMTQSSGSRASLPKV	326	
Dd	248 REFSEYLIQIYITPCMLIVTSWVSWFMDQGANPARVAVLGVTTLTLMATOSSINSALRPV	307	
OY	327 SYVKAIDIWMAVCCLLEVFALLLEVAALNEVSR--OHKEFIIRLRRO	370	
Dd	308 SYTKAIDVMTEGCVLTFVEGALLSESRFVNAYSASRDHNRNMKKARRR	353	
RESULT_9			
AAEI3312			
ID	AAEI3312 standard; Protein; 537 AA.		
XX			
AAEI3312:			
RNC			

```
XX 12-FEB-2002 (first entry)
DT Dermacentor variabilis clone 8 GABA-gated chloride channel protein.
XX
DE Gamma-aminobutyric acid; GABA; tick infestation; mite; antiparasitic;
KW GABA-gated chloride channel; recombinant expression; domestic animal.
XX
OS Dermacentor variabilis.
XX
PN WO200174884-A1.
XX
PD 11-OCT-2001.
XX
PE 28-MAR-2001; 2001WO-US09955.
XX
PR 31-MAR-2000; 2000US-193791P.
XX
PA (MERI ) MERCK & CO INC.
XX
PI Zheng Y, Cully D, Ludmerer S;
XX
DR WPI: 2002-010778/01.
DR N-PSDB: AAD22070.
XX
XX New polypeptide useful for preventing or treating tick infestation. In
PT humans, dogs, cattle, horses, deer, or other wild or domesticated
PT animals, comprises the Dermacentor variabilis gamma-aminobutyric acid
PT (GABA)-gated chloride channel.
XX
PS Claim 1: Fig 1; 59pp; English.
XX
CC The invention relates to gamma-aminobutyric acid (GABA)-gated chloride
CC channels and their corresponding nucleic acid molecules. GABA-gated
CC chloride channel proteins and DNA's are useful for preventing and
CC treating tick infestation, particularly in humans, dogs, cattle, horses,
CC deer, or other wild or domesticated animals. The nucleic acids are useful
CC as hybridisation probes or Polymerase Chain Reaction primers for
CC identifying the presence of Dermacentor variabilis GABA-gated chloride
CC channel nucleic acid or DNA's encoding for a GABA receptor. The nucleic
CC acids are also useful for the recombinant expression of D. variabilis
CC GABA-gated chloride channel proteins. GABA-gated chloride channel
CC proteins exert toxic effects on other ticks or related parasites such as
CC mites. The present sequence is Dermacentor variabilis clone 8 GABA-gated
CC chloride channel protein.
XX
SO Sequence 537 AA:
Query Match 32.7%; Score 736; DB 23; Length 537;
Best Local Similarity 40.9%; Pred. No. 1.4e-67;
Matches 158; Conservative 68; Mismatches 110; Indels 50; Gaps 9;
OY 55 TSGYDARIRPNFKGPVNVTCNIFNSFSYTKTMDYRVVFLRQOWNDRSLSYREYD 114
DB 52 TRGYDRRVRPVNGVPEVGTMOIISISTVSEVOMDPTSFYRROSBRDLRSFOKSPD 111
OY 115 -DSLDLPSMDSITWKPDLPFANEKGANFHEVTTDNKLRIFKGNVLYSIRLTLLISCL 173
DB 112 LESMTVGAEVAERIRWVDPTEFANEKSAFHAATPNTFLRIGSGGEVFRSIRLTVTAGCP 171
OY 174 MDLKNFMDIQTCTMQLSSSILCSPLPSLSLVGYTKMDLVFEMLEDAPAVOVAEGITL 233
DB 172 MDLRYFPMDROACQIEIE-----SFGYTKMDIRYKMSDQDTSVRIAKEVEL 217
OY 234 PQF-----ILRDEKDLGCTKRYNNGKFCIEVKFHELEOMGYLYIOMTIPSLIYISWV 289
DB 218 PQFVGLGHVQAKAEVALTGNYS--RLVC-EIRF--ARSMGYLYIOTIIPAGLIVISWV 272
OY 290 SFNIMMDAAPAVGIGITVTVTMTTOSGSRASLPKSVYKAIDIMAWVCLLFEAALE 349
DB 273 SFNHLRDSAPRAVALGVTVTTLMTTLMSTNAALPKISYKSIDVYLTCTCVMTFTALLE 332
OY 350 YAAINVEYSRQHKETRLRRORRORLEEDIIQESREYFRGYGLGHCLQARDGPGMEGSGI 409
```

```
DB 333 YAAVGYLCK-RITMRKTRCQQLAKLAHQHQR-----RCAASSNERSSEPL 378
OY 410 YSPQ-----PAPLRGCE 423
DB 379 ASPEVSIVKTVGSCVCPAAVASOGO 404
RESULT 10
AEI1314
ID AEI1314 standard; Protein: 537 AA.
XX
AC AEI1314;
XX
DT 12-FEB-2002 (first entry)
XX
DE Dermacentor variabilis clone 5 GABA-gated chloride channel protein.
XX
KW Gamma-aminobutyric acid; GABA; tick infestation; mite; antiparasitic;
KW GABA-gated chloride channel; recombinant expression; domestic animal.
XX
OS Dermacentor variabilis.
XX
PN WO200174884-A1.
XX
PD 11-OCT-2001.
XX
PE 28-MAR-2001; 2001WO-US09955.
XX
PR 31-MAR-2000; 2000US-193791P.
XX
PA (MERI ) MERCK & CO INC.
XX
PI Zheng Y, Cully D, Ludmerer S;
XX
DR WPI: 2002-010778/01.
DR N-PSDB: AAD22072.
XX
XX New polypeptide useful for preventing or treating tick infestation. In
PT humans, dogs, cattle, horses, deer, or other wild or domesticated
PT animals, comprises the Dermacentor variabilis gamma-aminobutyric acid
PT (GABA)-gated chloride channel.
XX
PS Claim 1: Fig 1; 59pp; English.
XX
CC The invention relates to gamma-aminobutyric acid (GABA)-gated chloride
CC channels and their corresponding nucleic acid molecules. GABA-gated
CC chloride channel proteins and DNA's are useful for preventing and
CC treating tick infestation, particularly in humans, dogs, cattle, horses,
CC deer, or other wild or domesticated animals. The nucleic acids are useful
CC as hybridisation probes or Polymerase Chain Reaction primers for
CC identifying the presence of Dermacentor variabilis GABA-gated chloride
CC channel nucleic acid or DNA's encoding for a GABA receptor. The nucleic
CC acids are also useful for the recombinant expression of D. variabilis
CC GABA-gated chloride channel proteins. GABA-gated chloride channel
CC proteins exert toxic effects on other ticks or related parasites such as
CC mites. The present sequence is Dermacentor variabilis clone 5 GABA-gated
CC chloride channel protein.
XX
SO Sequence 537 AA:
Query Match 32.5%; Score 732; DB 23; Length 537;
Best Local Similarity 40.7%; Pred. No. 3.6e-67;
Matches 157; Conservative 68; Mismatches 111; Indels 50; Gaps 9;
OY 55 TSGYDARIRPNFKGPVNVTCNIFNSFSYTKTMDYRVVFLRQOWNDRSLSYREYD 114
DB 52 TRGYDRRVRPVNGVPEVGTMOIISISTVSEVOMDPTSFYRROSBRDLRSFOKSPD 111
OY 115 -DSLDLPSMDSITWKPDLPFANEKGANFHEVTTDNKLRIFKGNVLYSIRLTLLISCL 173
DB 112 LESMTVGAEVAERIRWVDPTEFANEKSAFHAATPNTFLRIGSGGEVFRSIRLTVTAGCP 171
```



XX 28-MAR-2001; 2001WO-US09955.  
PF  
XX  
PR 31-MAR-2000; 2000US-193791P.  
XX  
PA (MERI ) MERCK & CO INC.  
XX  
PI Zheng Y, Cully D, Ludmerer S;  
XX  
DR WPI: 2002-010778/01.  
N-PSDB; AAD22071.  
XX  
PT New polypeptide useful for preventing or treating tick infestation, in  
PT humans, dogs, cattle, horses, deer, or other wild or domesticated  
PT animals, comprises the Dermacentor variabilis gamma-aminobutyric acid  
PT (GABA)-gated chloride channel  
XX  
PS Claim 1; Fig 1; 59pp; English.  
XX  
CC The invention relates to gamma-aminobutyric acid (GABA)-gated chloride  
CC channels and their corresponding nucleic acid molecules. GABA-gated  
CC chloride channel proteins and DNA's are useful for preventing and  
CC treating tick infestation, particularly in humans, dogs, cattle, horses,  
CC deer, or other wild or domesticated animals. The nucleic acids are useful  
CC as hybridisation probes or Polymerase Chain Reaction primers for  
CC identifying the presence of Dermacentor variabilis GABA-gated chloride  
CC channel nucleic acid or DNA's encoding for a GABA receptor. The nucleic  
CC acids are also useful for the recombinant expression of D. variabilis  
CC GABA-gated chloride channel proteins. GABA-gated chloride channel  
CC proteins exert toxic effects on other ticks or related parasites such as  
CC mites. The present sequence is Dermacentor variabilis clone 9 GABA-gated  
CC chloride channel protein.  
XX  
SO Sequence 537 AA:  
Query Match 32.5%; Score 731; DB 23; Length 537;  
Best Local Similarity 40.7%; Pred. No. 4.6e-67;  
Matches 157; Conservative 69; Mismatches 110; Indels 50; Gaps 9;  
QY 55 TSGYDARIRPNFGPVPVNTCNIFNSFSVTMTDYNVFLRQOWNDPRLSYREYP 114  
DB 52 TGTGDRKVRPNFGVGEVGVMTQIISIVSVQWDFSDFFKOSWDRERLSFQKSPD 111  
QY 115 -DSLDLDPMSLDSINRDLFFANKEGANFHEVTDNKLIRFKGNVLSIRLLISCL 173  
DB 112 LESMTVGAEVARIWPDFFANESKAYFHAATTPTFLRISGGEVFSIRLTVAACP 171  
QY 174 MDLKNFMDIQCTWQLESSSILCSPLSLISVGYTMKDLVEFEMLEDAVAOVAEGLT 233  
DB 172 MDLRYPMDROACTIEIE-----SFGYTMKDIRYMSDGDTSVRIAKEVEL 217  
QY 234 POF-----ILRDEKDLCCCKHYNTGKFTCTIEVKFHLEROMGYLLIOMYPSLLIVLSW 289  
DB 218 POKVLGHVOKAEVLLTGNYS--RLVC-EIRF--ARSGYLLIYIPAGLIVVSW 272  
QY 290 SFWINDDAARVAGLITTVLTMTTOSGSRASLPSVSYKAIDIMAAVCLLFFAALLE 349  
DB 273 SFPLHNASARVALGVTVLTMTTILMSSNALPISIVKSIDVLTGTCFVAVFALLE 332  
QY 350 YAAINVSROHKEFFILRRRORRORLEEDIIQESRFYFRGYGLHCLQARFGSPMEGSGT 409  
DB 333 YAAVGYLGR--RITMKRTKCOOLAKLAEOHRQ-----RCAAAASNPSSEPLL 378  
QY 410 YSPQ-----PAPLREGE 423  
DB 379 ASPEVSIKTVGSCOVCPAAVASOQ 404  
RESULT 13  
AAW05246  
ID AAW05246 standard; protein: 456 AA.  
XX  
AC AAW05246;

XX 09-FEB-1997 (first entry)  
DT  
XX  
DE Drosophila glutamate-gated chloride channel.  
XX  
KW glutamate gated chloride channel; GluCl; ligand gated channel;  
KW avermectin binding protein; glutamate binding protein;  
KW antiparasitic; anthelmintic; helminthiasis; ectoparasiticide;  
KW acaricide; insecticide; pesticide.  
XX  
OS Drosophila melanogaster strain Oregon R.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..22  
FT /label= Sig-peptide  
FT Protein 23..456  
FT /label= Mat-protein  
XX  
PN WO9634940-A1.  
XX  
PD 07-NOV-1996.  
XX  
PF 01-MAY-1996; 96WO-US06035.  
XX  
PR 05-MAY-1995; 95US-0435933.  
XX  
PA (MERI ) MERCK & CO INC.  
XX  
PI Arena JP, Cully DF, Liu KK, Pareess PS;  
XX  
DR WPI: 1996-506147/50.  
DR N-PSDB; AAT43596.  
XX  
PT New nucleic acid encoding Drosophila avermectin and/or glutamate  
PT binding protein - useful for isolating cpds. used as antiparasitic  
PT agents for the treatment or prevention of helminthiasis in domestic  
PT animals  
XX  
PS Claim 11; Page 45-46; 64pp; English.  
XX  
CC A novel Drosophila glutamate gated chloride channel (GluCl) (AAW05246)  
CC is selectively opened by either avermectin or glutamate and  
CC represents a target of avermectin action in arthropods. Its amino  
CC acid sequence was deduced from a Drosophila head cDNA clone (AAT43596).  
CC Recombinant GluCl, or host cells expressing it, can be used to screen  
CC for modulators of GluCl that can be used as ectoparasiticide,  
CC antiparasitic, anthelmintic, acaricide or insecticide agents for  
CC animals or plants.  
XX  
SO Sequence 456 AA:  
Query Match 32.4%; Score 730.5; DB 17; Length 456;  
Best Local Similarity 42.3%; Pred. No. 4e-67;  
Matches 154; Conservative 61; Mismatches 120; Indels 29; Gaps 7;  
QY 21 VLRLVALAKEEVSCTKGSOPMSPDFLKLGMRTSGYDARIRP---NFKGPVNTCNI 77  
DB 10 ILTFASLCSLSLNNNAKVNREKEKYLVDQILG-AKGYDARRPSGNGIDGAIYRINL 68  
QY 78 FINSFSSVTKTMTDYNVFLRQOWNDPRLSYREYPDSDLDLPSMLDSIMKRDLEFFANE 137  
DB 69 FVRSIMTISDIKMEYSVQLTFRQWMDERLKFDDIGRLKYLLTTEANRWMPDLEFSNE 128  
QY 138 KGANFHEVTDNKLIRFKGNVLSIRLLISCLMDLKNFMDIQCTWQLESSSILC 197  
DB 129 KEGFHNIIIMPVNYIRIFPNGSVLSIRISLTACPNKLILYPLDQICSLRN----- 181  
QY 198 SPLPSLSVGYTMKDLVEFEMLEDAVAOVAEGLTLPOFILLRDEKDL-GCCTKHYNTGKF 256  
DB 182 -----ASYGMVTNDLVFLMKRGDP-VGVYKKHLHLPFL--EKFLTDVCSNKTMTGEV 231  
QY 257 TCIEVKFHLEROMGYLLIOMYPSLLIVLSWSPWINDDAARVAGLITTVLTMTTOS 316







CC beneficial insect may render the insect resistant to a pesticide and can  
CC be used in combination with pesticides in the field to reduce or  
CC eliminate the presence of harmful insects only. The insect GABA receptor  
CC shows critical pharmacological differences from the vertebrate receptor  
CC which may result in the development of insect-specific insecticides with  
CC greater safety for vertebrate exposure. This sequence represents a  
CC variant *Drosophila simulans* (fruitfly) GABA receptor R1MD-R described  
CC in the method of the invention.  
XX

SQ Sequence 637 AA:

Query Match 32.28; Score 726; DB 21; Length 637;

Best Local Similarity 36.7%; Pred. No. 1.9e-66;

Matches 156; Conservative 86; Mismatches 143; Indels 40; Gaps 9;

```
OY 5 VPATLSFLLLMTLPGOVLNVALAKE-----VKSGTGSGSPMSDFLDKLMGRTS 56
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
Db 46 LPRT-PLTTW-----LAINMALIAQETGHRHHTVQAATGGGMLGDVNIASILDSESV 99
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
OY 57 GYDARIIRNFKPPVNVYTCNIFINSFSSVKTMDYRVNVELRQWNPRLSYREYPD-D 115
   || | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 100 SYDKRVNPNGGPPVEGVMTMYVLSVSEVLMDFLDFEFROFWTDPRLAIRKRPGVE 159
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
OY 116 SLDLPSMLDSIMKPDLEFANEGKANPEVTTDNKLLRIFKNGNVLYSIRLTLLSCAMD 175
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
Db 160 TLSVGESEFIKNIMWPDFFVNEKOSYFHIAFTSNEFIRVHSGSITRSIRLTITASCPMN 219
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
OY 176 LKNFPMIOCTQMOLESSILCSPLSLSVGTMKDLVFEULEDAPAVQVAGLTLPQ 235
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
Db 220 LQFFPMDRQLCHIEI-----SFGYTMRDIRFWRDGLSVGMSSEVELPQ 265
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
OY 236 FILDEKDLCCCTKHYNTGKFTCIKVEFHLEROMGYLLIOMYIPSLIIVLSWSEFWM 295
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 266 FRVLGHNQ-RATEINLTGMYSLACIQVRSKGYLLIQIYIPSGLIIVISWSEFWMNR 324
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
OY 296 DAAPARVGLITTVLINTTQSSGSRASLPVSVYKAIIDIMMAVCLLFVFAALLEYAINF 355
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
Db 325 NATPARVGLGVTVLTMTLMSSTNALPKISYKSIDVYLGTCFVWVFASLLEYATVGY 384
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
OY 356 VSRQHKFEIIRLRQRQRQRLIEDIIOESRYFRGYGHC---LQARDGPMESGIYSP 412
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
Db 385 MAK-----RIQMKQRKFMALQIAEQKKOOLDGANOQOANPNPNANVGGP-GGVGVPG 437
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
OY 413 QPPAP 417
   | |
Db 438 GPGGP 442
```

Search completed: June 25, 2003, 17:15:32  
Job time : 64.6734 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 25, 2003, 17:15:40 ; Search time 35.6756 Seconds  
(without alignments)  
1307.255 Million cell updates/sec

Title: US-10-075-846-4  
Perfect score: 2252  
Sequence: 1 MTLVLPATLSFLMLTLPCO.....POPPAPLRGETTRKLYVD 431

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 41779 seqs, 108206813 residues

Total number of hits satisfying chosen parameters: 41779

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications\_AA: \*  
1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep: \*  
2: /cgn2\_6/ptodata/1/pubpaa/BCI\_NEW\_PUB.pep: \*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep: \*  
4: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep: \*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep: \*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep: \*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep: \*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep: \*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep: \*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep: \*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep: \*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep: \*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep: \*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2252	100.0	431	US-10-075-846-4	Sequence 4, Appl1
2	2163	96.0	417	US-10-075-846-2	Sequence 2, Appl1
3	1682.5	74.7	452	US-10-075-846-13	Sequence 12, Appl1
4	1678	74.5	337	US-10-075-846-12	Sequence 13, Appl1
5	1638	72.7	312	US-10-075-846-15	Sequence 15, Appl1
6	1602.5	71.2	465	US-10-075-846-11	Sequence 11, Appl1
7	1580.5	70.2	449	US-10-075-846-10	Sequence 10, Appl1
8	1549	68.8	298	US-09-969-844-14	Sequence 14, Appl1
9	737	32.7	450	US-09-969-844-14	Sequence 14, Appl1
10	724	32.1	533	US-09-808-602-87	Sequence 87, Appl1
11	724	32.1	533	US-09-808-198-74	Sequence 74, Appl1
12	719	31.9	397	US-10-239-420-2	Sequence 2, Appl1
13	678.5	30.1	452	US-10-211-673-12	Sequence 12, Appl1
14	655	29.1	425	US-10-239-420-5	Sequence 5, Appl1
15	652.5	29.0	440	US-09-510-662A-39	Sequence 39, Appl1
16	652.5	29.0	440	US-09-778-320-39	Sequence 39, Appl1
17	652.5	29.0	440	US-09-910-689-39	Sequence 39, Appl1
18	652.5	29.0	440	US-10-010-742-39	Sequence 39, Appl1
19	652.5	29.0	461	US-10-106-698-6308	Sequence 6308, Appl1

20	632.5	28.1	397	US-09-808-483-6	Sequence 6, Appl1
21	632.5	28.1	422	US-09-808-483-4	Sequence 4, Appl1
22	632.5	28.1	423	US-09-808-483-8	Sequence 8, Appl1
23	620	27.5	465	US-09-818-657-2	Sequence 2, Appl1
24	620	27.5	465	US-09-818-657-2	Sequence 2, Appl1
25	619.5	27.5	554	US-10-211-673-8	Sequence 8, Appl1
26	618	27.4	465	US-09-818-657-4	Sequence 4, Appl1
27	615	27.3	464	US-09-839-446-35	Sequence 35, Appl1
28	615	27.3	464	US-09-839-446-35	Sequence 35, Appl1
29	615	27.3	464	US-09-839-446-37	Sequence 37, Appl1
30	615	27.3	464	US-09-839-446-38	Sequence 38, Appl1
31	615	27.3	464	US-09-742-311-4	Sequence 4, Appl1
32	615	27.3	464	US-09-898-570-35	Sequence 35, Appl1
33	615	27.3	464	US-09-898-570-36	Sequence 36, Appl1
34	615	27.3	464	US-09-898-570-37	Sequence 37, Appl1
35	615	27.3	464	US-09-898-570-38	Sequence 38, Appl1
36	615	27.3	464	US-09-808-483-10	Sequence 10, Appl1
37	615	27.3	535	US-09-808-483-12	Sequence 12, Appl1
38	608	27.0	467	US-09-765-069-2	Sequence 2, Appl1
39	608	27.0	467	US-09-742-311-2	Sequence 2, Appl1
40	607	27.0	420	US-09-765-069-8	Sequence 8, Appl1
41	603	26.8	392	US-09-765-069-4	Sequence 4, Appl1
42	602	26.7	345	US-09-765-069-10	Sequence 10, Appl1
43	594.5	26.4	468	US-09-839-446-6	Sequence 6, Appl1
44	594.5	26.4	468	US-09-898-570-6	Sequence 6, Appl1
45	592.5	26.3	466	US-09-839-446-8	Sequence 8, Appl1

ALIGNMENTS

RESULT 1  
US-10-075-846-4  
; Sequence 4, Application US/10075846  
; Publication No. US20030032608A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING A NOVEL GLYCINE RECEPTOR ALPHA SUBUNIT  
; FILE REFERENCE: D0079 NP  
; CURRENT APPLICATION NUMBER: US/10/075, 846  
; PRIOR FILING DATE: 2002-02-13  
; PRIOR APPLICATION NUMBER: US 60/269, 535  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 81  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 4  
; TYPE: PRT  
; LENGTH: 431  
; ORGANISM: homo sapiens  
US-10-075-846-4

Query Match 100.0%; Score 2252; DB 9; Length 431;  
Best Local Similarity 100.0%; Pred. No. 1; le-206;  
Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MTLVLPATLSFLMLTLPGOVLLRVALAKEVSGTGSGOPMSDFDLKMGRTSGYDA	60
DB	1	MTLVLPATLSFLMLTLPGOVLLRVALAKEVSGTGSGOPMSDFDLKMGRTSGYDA	60
QY	61	RIRNFKGPPNVNVCNFINFSFVYTKTMDYRVNVLFRQOMNDRSLYREYPPDSLDLD	120
DB	61	RIRNFKGPPNVNVCNFINFSFVYTKTMDYRVNVLFRQOMNDRSLYREYPPDSLDLD	120
QY	121	PSMLDSIWKPDLPFANEKGFHEVYTDNKLRIFFKNGVLYSTRITLILSCLMDLNKP	180
DB	121	PSMLDSIWKPDLPFANEKGFHEVYTDNKLRIFFKNGVLYSTRITLILSCLMDLNKP	180
QY	181	MDIOTCFMQLSSSTLSPLSLISVGYTKKDLVFEFLDPAPVNOVAEGTLTLPFILRD	240
DB	181	MDIOTCFMQLSSSTLSPLSLISVGYTKKDLVFEFLDPAPVNOVAEGTLTLPFILRD	240
QY	241	EKDGCCCTKHVNTGKFTCIEVKFHLEROMGYLLIOMYIPSLILVILSVSWFMINDAPA	300



ORGANISM: Mus musculus  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: (322)..  
OTHER INFORMATION: wherein "X" is any amino acid.  
US-10-075-846-12

Query Match 74.5%: Score 1678; DB 9; Length 337;  
Best Local Similarity 92.0%: Pred. No. 5,5e-152;  
Matches 323; Conservative 7; Mismatches 7; Indels 14; Gaps 1;

OY 25 VALAKEVKSSTKSGSPSPDFLDKLMGRTSGYDARIRPNFKGPPVAVTCNIFINSFS 84  
DB 1 VALAKEVKSSTKSGSPSPDFLDKLMGRTSGYDARIRPNFKGPPVAVTCNIFINSFS 60  
OY 85 VTKTMDYRVNVLROQNDPRLSYREYPPDSLDLPSMDSIMKPDLEFANEGANFHE 144  
DB 61 VTEETMDYRVNVLROQNDPRLSYREYPPDSLDLPSMDSIMKPDLEFANEGANFHE 120  
OY 145 VTTDNKLLRIFKNGNVLISIRLTILSLMDLKNFPMIDICTMOLESSILCSPLSLS 204  
DB 121 VTTDNKLLRIFKNGNVLISIRLTILSLMDLKNFPMIDICTMOLE----- 167  
OY 205 LSVGYTKDLVFELEDAVAQVAGLTPQFIRDEKDLCCCKHYNTGKFTCIEVKFH 264  
DB 168 -SEGYTNMDLFEWLEDAVAQVAGLTPQFIRDEKDLCCCKHYNTGKFTCIEVKFH 226  
OY 265 LEROMGYLLIOMYIPSLIIVLSWVSFMINMDAPARVGLGITTVMTTQSSGRASLP 324  
DB 227 LEROMGYLLIOMYIPSLIIVLSWVSFMINMDAPARVGLGITTVMTTQSSGRASLP 286  
OY 325 KVSIVKAIIDIMAVCLLEFVALLLEYAAINFVSROKHEFIRRRORRORL 375  
DB 287 KVSIVKAIIDIMAVCLLEFVALLLEYAAINFVSROKHEFIRRRORRORR 337

RESULT 5  
US-10-075-846-15  
Sequence 15, Application US/10075846  
Publication No. US20030032608A1  
GENERAL INFORMATION:  
APPLICANT: Bristol-Myers Squibb Company  
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING A NOVEL GLYCINE RECEPTOR ALPHA SUBUNIT  
FILE REFERENCE: D0079 NP  
CURRENT APPLICATION NUMBER: US/10/075, 846  
PRIOR FILING DATE: 2002-02-13  
PRIOR APPLICATION NUMBER: US 60/269, 535  
NUMBER OF SEQ ID NOS: 81  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 15  
LENGTH: 312  
TYPE: PRT  
ORGANISM: homo sapiens  
US-10-075-846-15

Query Match 72.7%: Score 1638; DB 9; Length 312;  
Best Local Similarity 100.0%: Pred. No. 3,3e-148;  
Matches 312; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 44 PSDFLDKLMGRTSGYDARIRPNFKGPPVAVTCNIFINSFSVTCTMDYRVNVLROQND 103  
DB 1 PSDFLDKLMGRTSGYDARIRPNFKGPPVAVTCNIFINSFSVTCTMDYRVNVLROQND 60  
OY 104 DPLRSYREYPPDSLDLPSMDSIMKPDLEFANEGANFHEVTTDNKLLRIFKNGNVLIS 163  
DB 61 DPLRSYREYPPDSLDLPSMDSIMKPDLEFANEGANFHEVTTDNKLLRIFKNGNVLIS 120  
OY 164 IRLTILISCLMDLKNFPMIDICTMOLESSILCSPLSLSLVGYTKDLVFELEDAV 223  
DB 121 IRLTILISCLMDLKNFPMIDICTMOLESSILCSPLSLSLVGYTKDLVFELEDAV 180

OY 224 AVQVAGLTPQFIRDEKDLCCCKHYNTGKFTCIEVKFHLEROMGYLLIOMYIPSLI 283  
DB 181 AVQVAGLTPQFIRDEKDLCCCKHYNTGKFTCIEVKFHLEROMGYLLIOMYIPSLI 240  
OY 284 VILSWVSFMINMDAPARVGLGITTVMTTQSSGRASLPKVSIVKAIIDIMAVCLLFV 343  
DB 241 VILSWVSFMINMDAPARVGLGITTVMTTQSSGRASLPKVSIVKAIIDIMAVCLLFV 300  
OY 344 FALLEVAIAINF 355  
DB 301 FALLEVAIAINF 312

RESULT 6  
US-10-075-846-11  
Sequence 11, Application US/10075846  
Publication No. US20030032608A1  
GENERAL INFORMATION:  
APPLICANT: Bristol-Myers Squibb Company  
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING A NOVEL GLYCINE RECEPTOR ALPHA SUBUNIT  
FILE REFERENCE: D0079 NP  
CURRENT APPLICATION NUMBER: US/10/075, 846  
PRIOR FILING DATE: 2002-02-13  
PRIOR APPLICATION NUMBER: US 60/269, 535  
NUMBER OF SEQ ID NOS: 81  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 11  
LENGTH: 465  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-075-846-11

Query Match 71.2%: Score 1602.5; DB 9; Length 465;  
Best Local Similarity 73.8%: Pred. No. 1,4e-144;  
Matches 312; Conservative 35; Mismatches 43; Indels 33; Gaps 6;

OY 22 LLRVALAKEEVKSGTKSGSPSPDFLDKLMGRTSGYDARIRPNFKGPPVAVTCNIFINS 81  
DB 21 LLSLVATKRTSDARSAPSPSPDFLDKLMGRTSGYDARIRPNFKGPPVAVTCNIFINS 80  
OY 82 FSSVYTKTMDYRVNVLROQNDPRLSYREYPPDSLDLPSMDSIMKPDLEFANEGAN 141  
DB 81 FGSIAETMDYRVNVLROQNDPRLSYREYPPDSLDLPSMDSIMKPDLEFANEGAN 140  
OY 142 FHEVTTDNKLLRIFKNGNVLISIRLTILSLMDLKNFPMIDICTMOLESSILCSPL 201  
DB 141 FHEVTTDNKLLRIFKNGNVLISIRLTILSLMDLKNFPMIDICTMOLE----- 190  
OY 202 SLISVGYTKDLVFELEDAVAQVAGLTPQFIRDEKDLCCCKHYNTGKFTCIEV 261  
DB 191 ----SPGTTMNDLFEWLEDAV-QVAGLTPQFIRDEKDLCCCKHYNTGKFTCIEV 245  
OY 262 KHLEROMGYLLIOMYIPSLIIVLSWVSFMINMDAPARVGLGITTVMTTQSSGRAS 321  
DB 246 RHLEROMGYLLIOMYIPSLIIVLSWVSFMINMDAPARVGLGITTVMTTQSSGRAS 305  
OY 322 SLRVSIVKAIIDIMAVCLLEFVALLLEYAAINFVSROKHEFIRRRORRORR-RL- 376  
DB 306 SLRVSIVKAIIDIMAVCLLEFVALLLEYAAINFVSROKHEFIRRRORRORR-RL- 365  
OY 377 -----EDIOESRFYFRGYGLCHLOARDGPMESGSIYSPQPPAPL-REGFTTKL 428  
DB 366 FYRSDMDDEVRESFSTAYGMGBCLOAKGOMPKG-----PNIHVQVMKSDDEMKV 420  
OY 429 YVD 431  
DB 421 FID 423

RESULT 7  
US-10-075-846-10

; Sequence 10, Application US/10075846  
; Publication No. US20030032608A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING A NOVEL GLYCINE RECEPTOR ALPHA SUBUNIT E  
; TITLE OF INVENTION: IN THE GASTROINTESTINAL TRACT, HGRA4, and SPLICE VARIANT THEREOF  
; FILE REFERENCE: D0079 NP  
; CURRENT APPLICATION NUMBER: US/10/075,846  
; CURRENT FILING DATE: 2002-02-13  
; PRIOR APPLICATION NUMBER: US 60/269,535  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 81  
; SOFTWARE: Patent version 3.0  
; SEQ ID NO 10  
; LENGTH: 449  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-075-846-10

Query Match 70.2% Score 1580.5; DB 9; Length 449;  
Best Local Similarity 73.4%; Pred. No. 1.7e-142;  
Matches 307; Conservative 38; Mismatches 50; Indels 23; Gaps 7;

QY 17 LPQVLLVALAKEVKSQSPMSDFLDKMGRTSGVDARIRNFKGPPVNVTCN 76  
DB 11 LSGAIVFSLASKAEARSAFTKPMSPSDFLDKMGRTSGVDARIRNFKGPPVNVSCN 70  
QY 77 IFINFSVVTKTMDYRVNVLRLQOMNDPRLSREYPPDSDLDPSMDSIMKPOLFFRAN 136  
DB 71 IFINFSIAETMTMYRNINFLRQOMNDPRLANRYPPDSDLDPSMDSIMKPOLFFRAN 130  
QY 137 EKGAFHEVTDNKLIRFKNGNVLSIRLLILSLMDLKNFPMDIOTCTMOLESSITL 196  
DB 131 EKGAFHEITDNKLIRISNGNVLSIRITLTACPMDLKNFPMDIOTCTMOLE----- 185  
QY 197 CSPPLSLSLSGYTKMDLVFEMLEDAPAVOAEGTLTPOFILRDKDGLCCCKHYNTGKF 256  
DB 186 -----SFGYTMNDLDFEMOEG-AYOYADGTLTPOFILRDKDGLCCCKHYNTGKF 235  
QY 257 TCIEVKFHLEROMGYLLIOMYIPSLIYLISVSWFMINMDAPARVGLITVLTMTQS 316  
DB 236 TCIEARFHLEROMGYLLIOMYIPSLIYLISVSWFMINMDAPARVGLITVLTMTQS 295  
QY 317 SGRASLKRVSIVKADIMAVCLFVFAALLEVAIAINFSQKHETRLRRORORLE 376  
DB 296 SGRASLKRVSIVKADIMAVCLFVFAALLEVAIAINFSQKHETRLRRORORLE 353  
QY 377 EDIIOSRFYFRGYLG-HCLQARDGPMGEG--SGIISPOPPARLLHGETTRKLY 430  
DB 354 EDIAGEGRNFSAYMGPAICLQAKDGISVKGANNSNTNP-PPAD-SKSPERMRLFT 409

RESULT 8  
US-10-075-846-14  
; Sequence 14, Application US/10075846  
; Publication No. US20030032608A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING A NOVEL GLYCINE RECEPTOR ALPHA SUBUNIT E  
; TITLE OF INVENTION: IN THE GASTROINTESTINAL TRACT, HGRA4, and SPLICE VARIANT THEREOF  
; FILE REFERENCE: D0079 NP  
; CURRENT APPLICATION NUMBER: US/10/075,846  
; CURRENT FILING DATE: 2002-02-13  
; PRIOR APPLICATION NUMBER: US 60/269,535  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 81  
; SOFTWARE: Patent version 3.0  
; SEQ ID NO 14  
; LENGTH: 298  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-10-075-846-14

Query Match 68.8% Score 1549; DB 9; Length 298;  
Best Local Similarity 95.2%; Pred. No. 9.7e-140;  
Matches 297; Conservative 0; Mismatches 1; Indels 14; Gaps 1;

QY 44 PSDFLDKMGRTSGVDARIRNFKGPPVNVTCNIFINFSVVTKTMDYRVNVLRLQOMN 103  
DB 1 PSDFLDKMGRTSGVDARIRNFKGPPVNVTCNIFINFSVVTKTMDYRVNVLRLQOMN 60  
QY 104 DPLRSREYPPDSDLDPSMDSIMKPOLFFANEGAFHEVTDNKLIRFKNGNVLS 163  
DB 61 DPLRSREYPPDSDLDPSMDSIMKPOLFFANEGAFHEVTDNKLIRFKNGNVLS 120  
QY 164 IRLTLILSLMDLKNFPMDIOTCTMOLESSITLSPPLSLSLSGYTKMDLVFEMLEDA 223  
DB 121 IRLTLILSLMDLKNFPMDIOTCTMOLE-----SFGYTKMDLVFEMLEDA 166  
QY 224 AVOVAEGTLTPOFILRDKDGLCCCKHYNTGKFTCIEVKFHLEROMGYLLIOMYIPSLI 283  
DB 167 AVOVAEGTLTPOFILRDKDGLCCCKHYNTGKFTCIEVKFHLEROMGYLLIOMYIPSLI 226  
QY 284 VILSVSWFMINMDAPARVGLITVLTMTQSSGRASLKRVSIVKADIMAVCLFV 343  
DB 227 VILSVSWFMINMDAPARVGLITVLTMTQSSGRASLKRVSIVKADIMAVCLFV 286  
QY 344 FAALLEVAIANF 355  
DB 287 FAALLEVAIANF 298

RESULT 9  
US-09-969-844-14  
; Sequence 14, Application US/09969844  
; Publication No. US20020192776A1  
; GENERAL INFORMATION:  
; APPLICANT: xiao-zhou Michelle Mang  
; APPLICANT: Xavier Georges Sarda  
; APPLICANT: Michael David Tomalski  
; APPLICANT: Vincent Paul Mary Wingate  
; TITLE OF INVENTION: Heliothis glutamate Receptor  
; FILE REFERENCE: A32815-1 072667.0178  
; CURRENT APPLICATION NUMBER: US/09/969,844  
; CURRENT FILING DATE: 2001-10-03  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: FASTSEQ for Windows Version 3.0  
; SEQ ID NO 14  
; LENGTH: 450  
; TYPE: PRT  
; ORGANISM: Heliothis virescens  
US-09-969-844-14

Query Match 32.7% Score 737; DB 9; Length 450;  
Best Local Similarity 45.7%; Pred. No. 6.7e-62;  
Matches 158; Conservative 53; Mismatches 111; Indels 24; Gaps 7;

QY 31 EYKSGTKSGSPMSDFLDKMGRTSGVDARIRP---NFKGPPVNVTCNIFINFSVVTK 87  
DB 26 ECMNGCKINFEKEKQIIDLIG--FGRYDARIRPSGINTDGPVAVSVINIFRSISKID 84  
QY 88 TTMDEVAVNVLRLQOMNDPRLSREYPPDSDLDPSMDSIMKPOLFFANEGAFHEVTT 147  
DB 85 VTMSVSVLTFRQWDLERLKNNGIKRYLTTLTEARVWMPDLFFSNEKEGHEHNTIM 144  
QY 148 DNKLIRFKNGNVLSIRLLILSLMDLKNFPMDIOTCTMOLESSITLSPPLSLSLSV 207  
DB 145 PNVIIRIFPNNGNVLSIRISLTLSCPMKLKLPDKQCSLRM-----ASY 190  
QY 208 GTYMKDLVFEWLEDAPOVAEGTLTPOFILRDKDL-GCCTKHNTGKFTCIEVKFHL 266  
DB 191 GWTDDLVFLMKEGDP-VQVVKNLILPFTLT--EKELTDYCNKNTNGEYGLKDLDFK 247  
QY 267 RQMGYYLLIOMYIPSLIYLISVSWFMINMDAPARVGLITVLTMTQSSGRASLKP 326  
DB 248 REFSYLLIOLYIPCCMLIVSVSWFMDQAVPARVLLGVTTLTMTATQSSGINASLPV 307







Job time : 37.6756 secs

```
QY 107 LSYREXPDSDL-DPSMLDSIMKPDLEFANEKGFHEVTTDNKLRIFKNGNVYSIR 165
      :      :      :      :      :      :      :      :      :      :
Db 110 MKNANL-TRSLDNDPRLKWKVPDYFPNAKHGEQFVTPNVLLRIYPTGDILYMR 168
QY 166 LTIILSCMDLKNFPMDIOTCTMOLESSILCSPLSLSVGYTKMDLVEWLEDAV 225
      :      :      :      :      :      :      :      :      :      :
Db 169 LKLFPSCKMMERYPDLRQVCSIELASF-----KTEVELOW-GNAEAV 213
QY 226 QVAGSLTLPQFIRDEKDCCTKHNTGKFCIEVKFHLEROMGYLLIOMYIPSLIYI 285
      :      :      :      :      :      :      :      :      :      :
Db 214 TMTSGKMAQFELQ-QSLTKSGAFOIGEYSCLRALNKRSGHHLVOSYLPSTLIYV 272
QY 286 LSNVSEFMINDAAPARVGLITVTLTWTQSSGRASLPKSYVKAIDIMAVCLLEFVA 345
      :      :      :      :      :      :      :      :      :      :
Db 273 VSNVSEFLVDALPARTLGLVTLTLTSSSDHQANLAPSYKALDVWMTCTMFEFA 332
QY 346 ALLEYAINFVSQ 359
      :      :      :      :      :      :      :      :      :      :
Db 333 AVLEFVSYLARR 346
```

## RESULT 15

```
US-09-510-662A-39
: Sequence 39, Application US/09510662A
: Patent No. US20020155125A1
: GENERAL INFORMATION:
: APPLICANT: Dillion, Davin C.
: APPLICANT: Wang, Craig H.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND DIAGNOSIS
: FILE REFERENCE: 210121.491C1
: CURRENT APPLICATION NUMBER: US/09/510.662A
: NUMBER OF SEQ. ID NOS: 41
: SOFTWARE: Patent In Ver. 2.0
: SEQ ID NO 39
: LENGTH: 440
: TYPE: PRF
: ORGANISM: Homo sapiens
US-09-510-662A-39
```

Query Match 29.0%; Score 652.5; DB 9; Length 440;

Best Local Similarity 39.1%; Pred. No. 7.6e-54;

Matches 133; Conservative 70; Mismatches 110; Indels 27; Gaps 7;

```
QY 55 TSGYDARIRPNFKGPPVNTCNIFINSFSSVTKTMDYRVNVLROQMDPRLSYREYD 114
      :      :      :      :      :      :      :      :      :      :
Db 45 TAGYNKFLRPNFGGEPOIALTLDIASISSISESMDYTATYLRQWMDQRLVFE--GN 102
QY 115 DSDLDPSMLDSIMKPDLEFANEKGFHEVTTDNKLRIFKNGNVYSIRLTIILSCM 174
      :      :      :      :      :      :      :      :      :      :
Db 103 KSFLLDLARLVEFLWNPDTYVESKSPFHEVTGNRLIRLESNGTVLYALRLITTVACNM 162
QY 175 DLKNFPMDIOTCTMOLESSILCSPLSLSVGYTKMDLVEWLEDAVPAVOVAEGTLT 234
      :      :      :      :      :      :      :      :      :      :
Db 163 DLSKYPMDTQCTKQLE-----SMGCDNDVETWLRGNDVSVRGLEHLRLA 208
QY 235 QFIRDEKDLGCGCTK-HYNTGKFTCIEVKFHLEROMGYLLIOMYIPSLIYIISWVSFWI 293
      :      :      :      :      :      :      :      :      :      :
Db 209 QYTI--ERYFTLYTRSOQETGNTRYLQFELRNRNVLYFILETVPSTPLVLSWVSFWI 266
QY 294 NMDAAPARVGLITVTLTWTQSSGRASLPKVS-VYKATDIMAVCLLEFVALLEYAA 352
      :      :      :      :      :      :      :      :      :      :
Db 267 SLDVSPARTICIGVTTVMTTLMIGSRTSLPNTNCFKALIDVYIGICFSFVGALLEYAV 326
QY 353 INFVSRQHKKEFIRLRORORLE---DIQESRFYFR 388
      :      :      :      :      :      :      :      :      :      :
Db 327 AHVSSIQQ---MAKKDGTGTRKEVEVSITNINSSISPF 363
```

Search completed: June 25, 2003, 17:18:50

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: June 25, 2003, 17:14:25 ; Search time 23.1409 Seconds  
(without alignments)  
548.002 Million cell updates/sec

Title: US-10-075-846-4  
Perfect score: 2252  
Sequence: 1 MTTLVPAFLSLFLMTLPGQ.....PQPPAPLLRGEETRKLYVD 431

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*

- 1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep:\*
- 2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep:\*
- 3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep:\*
- 4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep:\*
- 5: /cgn2\_6/ptodata/1/1aa/PCUTUS\_COMB.pep:\*
- 6: /cgn2\_6/ptodata/1/1aa/Backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	745	33.1	617	US-08-137-614A-24	Sequence 24, Appl
2	737	32.7	450	US-09-592-891A-14	Sequence 14, Appl
3	731.5	32.5	455	US-09-130-339-2	Sequence 2, Appl
4	730.5	32.4	456	US-08-435-933-6	Sequence 6, Appl
5	730.5	32.4	456	PCT-US96-06035-6	Sequence 6, Appl
6	730	32.4	423	US-09-627-650B-19	Sequence 19, Appl
7	726	32.2	637	US-08-072-064-8	Sequence 8, Appl
8	724.5	32.2	459	US-09-002-361-6	Sequence 6, Appl
9	724.5	32.2	467	US-09-002-361-3	Sequence 3, Appl
10	724.5	32.2	481	US-09-002-361-5	Sequence 5, Appl
11	724.5	32.2	496	US-09-002-361-2	Sequence 2, Appl
12	720	32.0	488	US-08-554-659-2	Sequence 2, Appl
13	720	32.0	488	US-08-554-659-4	Sequence 4, Appl
14	720	32.0	637	US-08-072-064-1	Sequence 1, Appl
15	720	32.0	637	US-08-072-064-4	Sequence 4, Appl
16	720	32.0	637	PCT-US92-08558-1	Sequence 1, Appl
17	718	31.9	637	US-08-072-064-6	Sequence 6, Appl
18	717	31.8	496	US-08-137-614A-2	Sequence 2, Appl
19	717	31.8	496	US-08-768-301-2	Sequence 2, Appl
20	714.5	31.7	383	US-09-627-650B-20	Sequence 20, Appl
21	712	31.6	474	US-08-417-330A-20	Sequence 20, Appl
22	705	31.3	617	US-08-137-614A-26	Sequence 26, Appl
23	691.5	30.7	474	US-08-459-100A-4	Sequence 4, Appl
24	683.5	30.4	487	US-09-627-650B-13	Sequence 13, Appl
25	683.5	30.4	487	US-09-436-063C-13	Sequence 13, Appl
26	680.5	30.2	617	US-08-137-614A-25	Sequence 25, Appl
27	678.5	30.1	452	US-08-809-802-8	Sequence 12, Appl

28	673	29.9	487	US-08-249-112-4	Sequence 4, Appl
29	673	29.9	487	PCT-US95-06556-4	Sequence 4, Appl
30	670.5	29.8	510	US-08-249-112-3	Sequence 3, Appl
31	670.5	29.8	510	PCT-US95-06556-3	Sequence 3, Appl
32	663	29.4	506	US-09-627-650B-15	Sequence 15, Appl
33	663	29.4	506	US-09-436-063C-15	Sequence 15, Appl
34	657	29.2	487	US-09-627-650B-14	Sequence 14, Appl
35	657	29.2	487	US-09-436-063C-14	Sequence 14, Appl
36	656	29.1	475	US-09-627-650B-16	Sequence 16, Appl
37	656	29.1	475	US-09-436-063C-16	Sequence 16, Appl
38	652.5	29.0	440	US-08-459-100A-2	Sequence 2, Appl
39	652.5	29.0	440	PCT-US94-09589-2	Sequence 2, Appl
40	644.5	28.6	453	US-08-417-330A-18	Sequence 18, Appl
41	628	27.9	686	US-08-768-301-4	Sequence 4, Appl
42	626	27.8	467	US-08-459-100A-3	Sequence 3, Appl
43	619.5	27.5	554	US-08-809-802-8	Sequence 8, Appl
44	618	27.4	451	US-08-417-330A-12	Sequence 12, Appl
45	610.5	27.1	492	US-08-417-330A-14	Sequence 14, Appl

## ALIGNMENTS

RESULT 1  
US-08-137-614A-24  
Sequence 24, Application US/08137614A  
Patent No. 5487976  
GENERAL INFORMATION:  
APPLICANT: Soderlund, David M.  
APPLICANT: Knipfle, Douglas C.  
APPLICANT: Henderson, Joseph E.  
TITLE OF INVENTION: Gene Encoding An Insect  
TITLE OF INVENTION: Gamma-Aminobutyric Acid (GABA) Receptor Subunit  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Nixon, Hargrave, Devans & Doyle  
STREET: Clinton Square, P.O. Box 1051  
CITY: Rochester  
STATE: New York  
COUNTRY: USA  
ZIP: 14603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/137,614A  
FILING DATE: 15-OCT-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Timilan, Susan J.  
REGISTRATION NUMBER: 34,103  
REFERENCE/DOCKET NUMBER: 19603/120  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (716)263-1636  
TELEFAX: (716)263-1600  
INFORMATION FOR SEQ ID NO. 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 617 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-137-614A-24

Query Match 33.1%, Score 745, DB 1, Length 617;

Best Local Similarity 42.2%; Pred. No. 2.9e+70;  
Matches 146; Conservative 71; Mismatches 95; Indels 34; Gaps 6;

QY 48 LDKLMGRSGVDARIRPFKGPVAVTCNIFNSSTYTKTTMDRVAVFLRQONNDRL 107  
DB 63 IDSLT--GGVDIRLPFSFGAPLEIGLIVLASPDISSEVMDVYTIMYINQVWDERL 119





```

: CURRENT FILING DATE: 2000-07-28
: PRIOR APPLICATION NUMBER: 09/436,063
: PRIOR FILING DATE: 1999-11-08
: PRIOR APPLICATION NUMBER: 60/107,727
: PRIOR FILING DATE: 1998-11-09
: NUMBER OF SEQ ID NOS: 50
: SOFTWARE: Patent In Ver. 2.1
: SEQ ID NO 19
: LENGTH: 423
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence:/note =
: OTHER INFORMATION: synthetic construct
: US-09-627-6508-19

Query Match          32.4%; Score 730; DB 4; Length 423;
Best Local Similarity 37.7%; Pred. No. 6.3e-69;
Matches 150; Conservative 87; Mismatches 125; Indels 36; Gaps 8;

OY 5 VPATLSFLMTLPGOVLRLAKKE-----VKSQTKGSPMSPSDFDLKMGRTS 56
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
Db 15 LPRT-PLTTIW-----LAINMALIAOETGHRKRIHTVOAATGGSMIGDYNISALIDFSV 68

OY 57 GYDARIRPNFGPPVNTCNIFINSFSSVTKTMDYRVNVPRLQOMNDPRLSYREPD-D 115
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
Db 69 SYDRVRPNRGPPVEGVNTVYVLSISSEVLMDFLDFYRFQFMTDPRLAYRKRPGVE 128

OY 116 SLDDPSMDSIMKPDLEFANEKGFHEVTTDKLRIKFNGLVYSIRLTLLSCLMD 175
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
Db 129 TLSGSEFIKINWPDFFEVNEKOSYFIATTSNEFIRVHSGSITRSIRLTIASCPMN 188

OY 176 LKNFPMDIQTCTMOLLESSILCSPLPSLSVGYTMKDLVEFEMLEDAPAVOVAGLTLPQ 235
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
Db 189 LQYPMRQOLCHIEI-----SFGYTMRDYRFMRDGLSSVMSSEVELPQ 234

OY 236 FILDEKDLGCTCTHYNTGKTCTEVKFLEROMGYLLIOMYITSLIYLISWSFMINM 295
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
Db 235 FRVLGHQ-RATEINLTGNTSLRACEIOFVRSKGYLLIOTIYISGLIVISWSFMINR 293

OY 296 DAAPARVGLITTVLTMTTSSGSRASLPKVSYKKAIDIMAWCLLPEFALLLEYAINF 355
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
Db 294 NATAPARVGLITTVLTMTTSSGSRASLPKVSYKKAIDIMAWCLLPEFALLLEYAINF 353

OY 356 VSR---QKKEFIRLR--RQRORLEEDIIQESRYF 387
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
Db 354 MAKRIOMKRFMAIOKIAEQKQOTPSDIDKYSRIYE 391

RESULT 7
US-08-072-064-8
: Sequence 8, Application US/08072064
: Patent No. 6008046
: GENERAL INFORMATION:
: APPLICANT: FRENCH-CONSTANT, RICHARD H.
: TITLE OF INVENTION: DRUG AND PESTICIDE SCREENING
: NUMBER OF SEQUENCES: 20
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: PETER G. CARROLL
: STREET: 220 Montgomery Street, Suite 2200
: CITY: San Francisco
: STATE: California
: COUNTRY: United States of America
: ZIP: 94104
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/072,064
: FILING DATE: 19930602
```

```

: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 770,881
: FILING DATE: 04-OCT-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: CARROLL, PETER G.
: REGISTRATION NUMBER: 32,837
: REFERENCE/DOCKET NUMBER: OPHD-00574
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415/705-8410
: TELEFAX: 415/397-8338
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 637 amino acids
: TYPE: AMINO ACID
: TOPOLOGY: unknown
: MOLECULE TYPE: peptide
: US-08-072-064-8

Query Match          32.2%; Score 726; DB 3; Length 637;
Best Local Similarity 36.7%; Pred. No. 3.2e-68;
Matches 156; Conservative 86; Mismatches 143; Indels 40; Gaps 9;

OY 5 VPATLSFLMTLPGOVLRLAKKE-----VKSQTKGSPMSPSDFDLKMGRTS 56
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
Db 46 LPRT-PLTTIW-----LAINMALIAOETGHRKRIHTVOAATGGSMIGDYNISALIDFSV 99

OY 57 GYDARIRPNFGPPVNTCNIFINSFSSVTKTMDYRVNVPRLQOMNDPRLSYREPD-D 115
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
Db 100 SYDRVRPNRGPPVEGVNTVYVLSISSEVLMDFLDFYRFQFMTDPRLAYRKRPGVE 159

OY 116 SLDDPSMDSIMKPDLEFANEKGFHEVTTDKLRIKFNGLVYSIRLTLLSCLMD 175
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
Db 160 TLSGSEFIKINWPDFFEVNEKOSYFIATTSNEFIRVHSGSITRSIRLTIASCPMN 219

OY 176 LKNFPMDIQTCTMOLLESSILCSPLPSLSVGYTMKDLVEFEMLEDAPAVOVAGLTLPQ 235
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
Db 220 LQYPMRQOLCHIEI-----SFGYTMRDYRFMRDGLSSVMSSEVELPQ 265

OY 236 FILDEKDLGCTCTHYNTGKTCTEVKFLEROMGYLLIOMYITSLIYLISWSFMINM 295
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
Db 266 FRVLGHQ-RATEINLTGNTSLRACEIOFVRSKGYLLIOTIYISGLIVISWSFMINR 324

OY 296 DAAPARVGLITTVLTMTTSSGSRASLPKVSYKKAIDIMAWCLLPEFALLLEYAINF 355
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
Db 325 NATAPARVGLITTVLTMTTSSGSRASLPKVSYKKAIDIMAWCLLPEFALLLEYAINF 384

OY 356 VSRQKKEFIRLRQRORLEEDIIQESRYFRTYGHC---LQARDGGMBSGGLYSP 412
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
Db 385 MAK-----RIOMKRFMAIOKIAEQKQOOLDGANOQANPNPNANVGPP-GGVGVGPG 437

OY 413 QPPAP 417
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
Db 438 GPGCP 442

RESULT 8
US-09-002-361-6
: Sequence 6, Application US/09002361
: Patent No. 6329516
: GENERAL INFORMATION:
: APPLICANT: Halling, Blalk
: TITLE OF INVENTION: Lepidopteran GABA-gated Chloride
: NUMBER OF SEQUENCES: 43
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Dechert Price & Rhoads
: STREET: 997 Lenox Drive, Building 3, Suite 210
: CITY: Lawrenceville
: STATE: NJ
: COUNTRY: USA
: ZIP: 08543
: COMPUTER READABLE FORM:
```



```

CORRESPONDENCE ADDRESS:
ADDRESS: Dechert Price & Rhoads
STREET: 997 Lenox Drive, Building 3, Suite 210
CITY: Lawrenceville
STATE: NJ
COUNTRY: USA
ZIP: 08543
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/002.361
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
ATTORNEY/AGENT INFORMATION:
NAME: Bloom, Allen
REGISTRATION NUMBER: 29,135
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 609-520-3214
TELEFAX: 609-520-3259
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 481 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: Internal
US-09-002-361-5

Query Match      32.2%; Score 724.5; DB 4; Length 481;
Best Local Similarity 37.8%; Pred. No. 3e-68;
Matches 153; Conservative 79; Mismatches 120; Indels 53; Gaps 8;

58 YDARIRPNKGGPPVAVTNCIIFINSFSSVTKTMDRYVNFLEQONDPLSYREYD-DS 116
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
47 YDKRRPNYGGPPVEGYTMVLSISYSVYMDFTLDYFRQFTDPLATKKTGYET 106
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
117 LDLDPSMDSIMKPDLEFANEKGANFHEVTTDNKLRIFKNGNVYSIRLTLILSLMDL 176
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
107 LSVGESEFKINIWVPTFFVNEKQSFHATISNEFIRIHSSGSTRSIRLTITASCPLML 166
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
177 KNFPDIOTCTMOLESSSILCSPLPSLSISVGYTKDLVEFWLEDAPOVAEGTLTPOF 236
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
167 QYFPMRDQLCHIEI-----SFGYTRDIRYKWNENGPNSGVSEVSLPOF 212
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
227 ILRDEKDG---CCTKHYNTEKFTCIKFKHLEKQMGYLLQMTIPSLIYILSMVSW 292
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
213 -----KVLGHORAMEISLTGNYSLRACEIOFVSMGYLLQIYIPSLIYILSMVSW 267
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
293 INMDAPARVGLGITVTLMTQSSGSRASLPKVSYVKAIDIMAVCLLEFPAALLEYAA 352
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
268 LNRNATPARVALGVTTLMTLMSSTNALPKISYVSIDVYLGTCFPMVFASLLEYAT 327
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
333 INFSVR---QKKEFIRLRQRQRORLEEDT--IQESRFYRGVGLGHCLQAR----- 399
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
328 VGYMAKRIQMRKQRFVAIQIKASEKKIPVDCPPVGDPTLSKMGTLGRCPGRPSEVRPK 387
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
400 -----DGGPME-----GSGIYSPQAPAPILRGEETIRKL 428
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
368 VHDPAHSGGTLTENTINGRSGAEENENGPPIHLHPEKDISKL 432
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

```

```

Patent No. 6329516
GENERAL INFORMATION:
APPLICANT: Halling, Blaik
TITLE OF INVENTION: Lepidopteran GABA-gated Chloride
NUMBER OF INVENTIONS: 43
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESS: Dechert Price & Rhoads
STREET: 997 Lenox Drive, Building 3, Suite 210
CITY: Lawrenceville
STATE: NJ
COUNTRY: USA
ZIP: 08543
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/002.361
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
ATTORNEY/AGENT INFORMATION:
NAME: Bloom, Allen
REGISTRATION NUMBER: 29,135
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 609-520-3214
TELEFAX: 609-520-3259
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 496 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: Internal
US-09-002-361-2

Query Match      32.2%; Score 724.5; DB 4; Length 496;
Best Local Similarity 39.6%; Pred. No. 3.1e-68;
Matches 152; Conservative 72; Mismatches 109; Indels 51; Gaps 8;

45 SDFDKMKRTSGYDARIRPNKGGPPVAVTNCIIFINSFSSVTKTMDRYVNFLEQOND 104
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
44 SAIDSL--SVYDKRVRPNYGGPPVDGYTMVLSISLSEVRMDFTLDYFRQFTD 100
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
105 PLSYREYD--SLDLDPSMDSIMKPDLEFANEKGANFHEVTTDNKLRIFKNGNVYS 163
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
101 PLATKKTGYETLSVGESEFKINIWVPTFFVNEKQSFHATISNEFIRIHSSGSTRS 160
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
164 IRLTLILSLMDLKNFPMDIOTCTMOLESSSILCSPLPSLSISVGYTKDLVEFWLEDA 223
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
161 IRLTLITASCPMDLQYFPMRDQLCHIEI-----SFGYTRDIRYKWNENGP 206
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
224 AVQVAEGTLTPOFILRDEKDG---CCTKHYNTEKFTCIKFKHLEKQMGYLLQMTIP 279
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
207 SVGSSEVSLPOF-----KVLGHORAMEISLTGNYSLRACEIOFVSMGYLLQIYIP 261
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
280 SLIYILSMVSWINMDAPARVGLGITVTLMTQSSGSRASLPKVSYVKAIDIMAVC 339
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
262 SGLIYIISVSWFLNRNATPARVSGVTTLMTLMSSTNALPKISYVSIDVYLGTG 321
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
340 LLEFPAALLEYAAINFSVROKKEFIRLRQRQRORLEEDTIOESRFYRGVGLGHCLQAR 399
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
322 FVMVFASLLEYATVGYMAK-----RIQMRKQRFVAIQIKASEKKIQI----- 363
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
400 DGGPMEGSGIYSP-----QPPAP 417
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

```



```

Db          364 DGPGBAEP1PPRTSTLSRRPPP 387

RESULT 12
US-08-554-659-2
: Sequence 2, Application US/08554659
: Patent No. 5767261
GENERAL INFORMATION:
APPLICANT: Mingale, Vincent
APPLICANT: Wolff, Mark
TITLE OF INVENTION: LEPIDOPTERAN GABA GATED CHLORIDE CHANNEL
TITLE OF INVENTION: AND METHODS OF USE THEREOF
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: USA
ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/554,659
FILING DATE:
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9732
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 488 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-554-659-2

Query Match      32.0%; Score 720; DB 1; Length 488;
Best Local Similarity 38.0%; Pred. No. 9,1e-68;
Matches 156; Conservative 75; Mismatches 106; Indels 74; Gaps 10;

45 SDPLDKLGIRSGYDARIPNPKGPPVAVTGNIFINSFSVTAKTM DYRVAVVELRQKND 104
+ + | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
44 SAIDSL---SVSYDKRKRPYNGCPRYDVGVNMVYLSSLSSEVMADTLDFYEQFWTD 100
| : : : | : : | : : | : : | : : | : : | : : | : : | : : |
105 PRLSYREYPD--DSLDDLPDSMLDSIWKPDLPFEANEKANFHEVTYDNKLRIEFGKNVLVS 163
| : : : | : : | : : | : : | : : | : : | : : | : : | : : |
101 PRLAYKKRGSTGEYETLSVGSEEFTRINIIVPDTFVNENKOSYFIATTSNEFIRIHSSGISRS 160
| : : : | : : | : : | : : | : : | : : | : : | : : | : : |
164 IRTLILSLMDLNKPMPDIQTCTMQLESSSLCSPLSLSISVGYTKKDLYFEWLEDPAR 223
| : : : | : : | : : | : : | : : | : : | : : | : : | : : |
161 IRLITTAACPMDLQTFPPMDROLCNIIE-----SFQTYMRDRIRKYNEGPN 206
| : : : | : : | : : | : : | : : | : : | : : | : : | : : |
224 AVOVAEGTLTPQFLIRDEKDLG---CCTKHNTGKFPTCIYEKKFPLEBQMGGYLLQMYIP 279
| : : : | : : | : : | : : | : : | : : | : : | : : | : : |
207 SVGVASEVSLPQF-----KVIGHQGRAMEISLTITGSSRASLCEIOFVGSMGVIQLQIYIP 261
| : : : | : : | : : | : : | : : | : : | : : | : : | : : |
260 SLIIYILSNVSWRMIMDAAPPARVGLGITTVYLTMTTQSSGRASLPEKSVYKAIDIMAAVC 339
| : : : | : : | : : | : : | : : | : : | : : | : : | : : |
262 SGLIYIISWSFWMELRNATPARVSLGVTTVYLTMTTLMSTNAALPKISVYSIDVYGTC 321
| : : : | : : | : : | : : | : : | : : | : : | : : | : : |
340 LLEFVAALLEFAAINFVSROHKEFILRLRROR---ORLED----- 378
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
322 FVMVFETSLLEATLVGTMAK-----RIOMRKORFTPAVDVKMOLDGPGSAEPIPPRTSTL 375

```

```
.QY      379 -----IIQEERF-----YFRYGIGHCHLQADGGMESGIYSPPPA 416
           : | | | | | : | : | : | : | : | | | |
Db       376 SRPPPSRLSEFRFYVHDPKAYSKGTLENTINGARGOP-----GPAPPA 419

RESULT 13
US-08-554-659-4
; Sequence 4, Application US/08554659
; Patent No. 5767261
; GENERAL INFORMATION:
; APPLICANT: Wolff, Mark
; TITLE OF INVENTION: LEPIDOPTERAN GABA GATED CHLORIDE CHANNEL
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: USA
; ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08554.659
FILING DATE:
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Digilio, Frank S
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9732
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEX: (516) 742-4366
TELEFAX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO.: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 488 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-554-659-4

Query Match          32.0%; Score 720; DB 1; Length 488;
Best Local Similarity 38.0%; Pred. No. 9,1e-68;
Matches 156; Conservative 75; Mismatches 106; Indels 74; Gaps 10.
```

Db 322 FVWFTSLLEYATVGYMAK-----RQMRKQRFYVAKQMDIDPGSGAEPIPPRSTL 375

Qy 379 -----IIQESRF-----YFRGYGLCHCQARDGPGMEGSGIYSP 416

Db 376 SRPPSRLSYFRFVHDPKAKSKGTLENTINGARGP-----GPA 419

RESULT 14

US-08-072-064-1

Sequence 1, Application US/08072064

Patent No. 6008046

GENERAL INFORMATION:

APPLICANT: FRENCH-CONSTANT, RICHARD H.

APPLICANT: JACKSON, MEYER B.

TITLE OF INVENTION: DRUG AND PESTICIDE SCREENING

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESSEE: PETER G. CARROLL

STREET: 220 Montgomery Street, Suite 2200

CITY: San Francisco

STATE: California

COUNTRY: United States of America

ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/072.064

FILING DATE: 19930602

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 770,881

FILING DATE: 04-OCT-1991

ATTORNEY/AGENT INFORMATION:

NAME: CARROLL, PETER G.

REGISTRATION NUMBER: 32,837

REFERENCE/DOCKET NUMBER: OPHD-00574

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415/705-8410

TELEFAX: 415/397-8338

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 637 amino acids

TYPE: AMINO ACID

TOPOLOGY: unknown

MOLECULE TYPE: peptide

ORGANISM: Drosophila melanogaster

POSITION IN GENOME:

CHROMOSOME/SEGMENT: III; polytene subregion 66F

MAP POSITION: approximately map unit 26

US-08-072-064-1

Query Match 32.0%; Score 720; DB 3; Length 637;

Best Local Similarity 36.5%; Pred. No. 1.4e-67;

Matches 155; Conservative 86; Mismatches 144; Indels 40; Gaps 9;

Qy 5 VPATLSFLMLTLPQGYLLRVALAKEE-----VKSQTKSGSPMSPSPDLKMGRTS 56

Db 46 LPRT-PLTTW-----LAIMMALIAQETGKRIHTVGAATGGGSLDVAISALIDFSV 99

Qy 57 GYDARIRPNKGPVNVTCNFIINSFSSVTKTMDRVNVFLRQWMDPRLSYREPD-D 115

Db 100 SYDKRVARNYGGPVEVGVTMYVLSISVSEVLMDFLDFYFQFWTDPRLAYRKRPGVE 159

Qy 116 SLIDIPSLDLSIMKPDLEFANEKANHEVTTDKLIRFKNGNVLYSIRLTLLSCIMD 175

Db 160 TLVSGSEFIKNIMVDPFFVNEKOSYPHIATTSNEFIRVHNSGSITRSIRLTITASCPMN 219

Qy 176 LKNRPMIOCTMQLJESSLISPLSLSVGTYTMDLVFEMLEDRANOVAGELTLPQ 235

Db 220 LQYFPMRDLQCHIEE-----SFQYTMDIRYFMRDGLSGVMSSEVELQ 265

Qy 236 FILDENDDLGCTTHYNTGKTCIEVKEFLERQGYLLQMYISLILVLSNVSFINM 295

Db 266 FVILGHNO-RATEINLTGNYSRLACEQFVRSNGYLLQIYISGLIYIVSWSFMLNR 324

Qy 296 DAAPARVGLITVLTMTQSSGSRASLPKVSYKAIIDIMAVCLLFPAALLEYAINE 355

Db 325 NATPARVALGTYTLTMTTLMSSSTNALPKISYKSIDVYLGTCFVAVFASLEYATVGY 384

Qy 356 VSRQKHEFIRLRQRORQLEEDIIQESRFYRGYGLCHC---LQARDGPGMEGSGIYSP 412

Db 385 MAX-----RQMRKQRFYVAKQMDIDPGSGAEPIPPRSTL 419

Qy 413 GPAP 417

Db 438 GPGP 442

RESULT 15

US-08-072-064-4

Sequence 4, Application US/08072064

Patent No. 6008046

GENERAL INFORMATION:

APPLICANT: FRENCH-CONSTANT, RICHARD H.

APPLICANT: JACKSON, MEYER B.

TITLE OF INVENTION: DRUG AND PESTICIDE SCREENING

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESSEE: PETER G. CARROLL

STREET: 220 Montgomery Street, Suite 2200

CITY: San Francisco

STATE: California

COUNTRY: United States of America

ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/072.064

FILING DATE: 19930602

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 770,881

FILING DATE: 04-OCT-1991

ATTORNEY/AGENT INFORMATION:

NAME: CARROLL, PETER G.

REGISTRATION NUMBER: 32,837

REFERENCE/DOCKET NUMBER: OPHD-00574

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415/705-8410

TELEFAX: 415/397-8338

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 637 amino acids

TYPE: AMINO ACID

TOPOLOGY: unknown

MOLECULE TYPE: peptide

US-08-072-064-4

Query Match 32.0%; Score 720; DB 3; Length 637;

Best Local Similarity 36.5%; Pred. No. 1.4e-67;

Matches 155; Conservative 86; Mismatches 144; Indels 40; Gaps 9;

Qy 5 VPATLSFLMLTLPQGYLLRVALAKEE-----VKSQTKSGSPMSPSPDLKMGRTS 56

Db 46 LPRT-PLTTW-----LAIMMALIAQETGKRIHTVGAATGGGSLDVAISALIDFSV 99

Qy 57 GYDARIRPNKGPVNVTCNFIINSFSSVTKTMDRVNVFLRQWMDPRLSYREPD-D 115

Db 100 SYDKRVARNYGGPVEVGVTMYVLSISVSEVLMDFLDFYFQFWTDPRLAYRKRPGVE 159



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 25, 2003, 17:13:55 ; Search time 28.9262 Seconds

(without alignments)  
1432.403 Million cell updates/sec

Title: US-10-075-846-4

Perfect score: 2252

Sequence: 1 MTTLVPAFLSFLMLTLPLGQ.....PPPPAPLLRGGCTTRKLYVD 431

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR\_73:\*  
2: PIR1:\*  
3: PIR2:\*  
4: PIR3:\*  
5: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1687	74.9	337	2	A49970 glycine receptor a
2	1682.5	74.7	452	2	S12381 glycine receptor a
3	1677	74.5	452	2	S14816 glycine receptor a
4	1671	74.2	452	2	S18836 glycine receptor a
5	1656.5	73.6	451	2	B49970 glycine receptor a
6	1592	70.7	464	2	A23682 glycine receptor a
7	1587.5	70.5	451	2	C49970 glycine receptor a
8	1580.5	70.2	459	2	S12382 glycine receptor a
9	1568.5	69.6	457	2	S20662 glycine receptor a
10	871	38.7	496	2	S46459 glycine receptor b
11	863.5	38.3	496	2	JH0165 glycine receptor b
12	855.5	38.0	497	2	G02031 glycine receptor b
13	742	32.9	499	2	S17785 gamma-aminobutyric
14	726	32.2	488	2	JH0359 gamma-aminobutyric
15	724	32.1	533	2	S33744 cyclodiene insecti
16	717	31.8	496	2	JH0603 gamma-aminobutyric
17	716	31.8	491	2	JH0828 gamma-aminobutyric
18	712	31.6	474	2	S04465 gamma-aminobutyric
19	712	31.6	474	2	S53531 gamma-aminobutyric
20	712	31.6	512	2	I53656 gamma-aminobutyric
21	706	31.3	474	2	JH0829 gamma-aminobutyric
22	705.5	31.3	1106	2	T25065 hypothetical prote
23	700.5	31.1	478	2	T24805 gamma-aminobutyric
24	697	31.0	474	2	S53530 gamma-aminobutyric
25	696	30.9	473	2	A55275 gamma-aminobutyric
26	695.5	30.9	474	2	B60039 gamma-aminobutyric
27	694.5	30.8	474	2	B27142 gamma-aminobutyric
28	692.5	30.8	474	2	A40336 gamma-aminobutyric
29	692.5	30.8	654	2	C87791 protein B0207.12 f

30	692	30.7	473	2	S53532 gamma-aminobutyric
31	690	30.6	473	2	S04466 gamma-aminobutyric
32	687.5	30.5	476	2	S11440 gamma-aminobutyric
33	687	30.5	437	2	S34469 gamma-aminobutyric
34	673	29.9	434	2	S50865 gamma-aminobutyric
35	670.5	29.8	461	2	S50864 gamma-aminobutyric
36	668	29.7	606	2	A41145 gamma-aminobutyric
37	664.5	29.5	449	2	A34625 gamma-aminobutyric
38	653.5	29.0	449	2	A36303 gamma-aminobutyric
39	653.5	28.0	495	2	T20754 hypothetical prote
40	648.5	28.8	449	2	I2630 GABA receptor del
41	646.5	28.7	465	2	A38079 gamma-aminobutyric
42	643.5	28.6	545	2	T27614 hypothetical prote
43	634.5	28.2	453	2	S11087 gamma-aminobutyric
44	634	28.2	474	2	S13086 gamma-aminobutyric
45	631.5	28.0	473	2	A38627 gamma-aminobutyric

#### ALIGNMENTS

```
RESULT 1
A49970
glycine receptor alpha-4 chain - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Nov-1995 #sequence_revision 17-Nov-1995 #text_change 20-Aug-1999
C:Accession: A49970
R:Matzenbach, B.; Maulet, Y.; Sefton, L.; Courtlier, B.; Ayver, P.; Guenet, J.L.; Betz
J. Biol. Chem. 269, 2607-2613, 1994
A:Title: Structural analysis of mouse glycine receptor alpha subunit genes. Identific
A:Reference number: A49970; MUID:94132024; PMID:7507926
A:Accession: A49970
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-337 <MAT>
A:Cross-references: GB:X75850; NID:9435513; PIDN:CAA53468.1; PID:9817957
C:Genetics:
A:Gene: Glad
C:Superfamily: acetylcholine receptor; transmembrane protein
C:Keywords: neurotransmitter receptor; protein

Query Match      74.9%  Score 1687;  DB 2;  Length 337;
Best Local Similarity 92.3%;  Pred. No. 3.7e-137;
Matches 324;  Conservative 7;  Mismatches 6;  Indels 14;  Gaps 1;

QY      25 VALAKEEYKSGTSGSPSPDLKMGRTSGYDARIRPNKGPVAVTCIFINSS 84
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      1 VALAKEDYKSGLGSPSPDLKMGRTSGYDARIRPNKGPVAVTCIFINSS 60
QY      85 VTETMDYRVNVFLRQNNDDPLSYREYPPDOSLDDPMSLDSIMKPDLEFANEKANFHE 144
      ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      61 VTETMDYRVNVFLRQNNDDPLSYREYPPDOSLDDPMSLDSIMKPDLEFANEKANFHE 120
QY      145 VTETMDYRVNVFLRQNNDDPLSYREYPPDOSLDDPMSLDSIMKPDLEFANEKANFHE 144
      ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      121 VTETMDYRVNVFLRQNNDDPLSYREYPPDOSLDDPMSLDSIMKPDLEFANEKANFHE 120
QY      145 VTETMDYRVNVFLRQNNDDPLSYREYPPDOSLDDPMSLDSIMKPDLEFANEKANFHE 144
      ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      121 VTETMDYRVNVFLRQNNDDPLSYREYPPDOSLDDPMSLDSIMKPDLEFANEKANFHE 120
QY      205 LSVGYTMDYRVNVFLRQNNDDPLSYREYPPDOSLDDPMSLDSIMKPDLEFANEKANFHE 264
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      168 LSVGYTMDYRVNVFLRQNNDDPLSYREYPPDOSLDDPMSLDSIMKPDLEFANEKANFHE 226
QY      265 LEROMGYLLIOMYIFSLILVILSNVSWFIMNDAPARVGLGITYVLVTGTOSSGSRASLP 324
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      227 LEROMGYLLIOMYIFSLILVILSNVSWFIMNDAPARVGLGITYVLVTGTOSSGSRASLP 286
QY      325 KVSYVKAIDIMWAVCLLEFVFAALLEYAAINFEVSRHKEFIRLRORROR 375
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      287 KVSYVKAIDIMWAVCLLEFVFAALLEYAAINFEVSRHKEFIRLRORROR 337
RESULT 2
S12381
glycine receptor alpha-2 chain - human
```

C:Species: Homo sapiens (man)  
 C:Date: 21-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 20-Aug-1999  
 C:Accession: S12381  
 R:Greeningloh, G.; Schmieden, V.; Schofield, P.R.; Seeburg, P.H.; Siddique, T.; Mohandas  
 EMBO J. 9, 771-776, 1990  
 A:Title: Alpha subunit variants of the human glycine receptor: primary structures, funct  
 A:Reference number: S12381; MUID:90183975; PMID:2155780  
 A:Accession: S12381  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-452 <GRE>  
 A:Cross-references: GB:X52008; NID:931848; PIDN:CAA36251.1; PID:931849  
 A:Note: the sequence from Fig. 2 is inconsistent with that from Fig. 1 in lacking 16-Phe  
 C:Superfamily: acetylcholine receptor  
 C:Keywords: transmembrane protein

Query Match 74.7%; Score 1682.5; DB:2; Length 452;  
 Best Local Similarity 79.0%; Pred. No. 1.3e-136;  
 Matches 324; Conservative 29; Mismatches 36; Indels 21; Gaps 5;

24 RVALAKE-EVKSQTRGSPSPDFLDKLMGRTSGYDARIRNFKGPPVNTCNFINSF 82  
 23 RFAFCQKDHDSRGKQPSQTLSPSDFLDKLMGRTSGYDARIRNFKGPPVNTCNFINSF 82  
 83 SSVTKTMDYRVNVEFLRQOMDPRLSREYDDSLDLPMSDLSIMKPDLEFANEKANF 142  
 83 GSVTETMDYRVNVEFLRQOMDPRLSREYDDSLDLPMSDLSIMKPDLEFANEKANF 142  
 143 HEVTDNKLRFKNGNVLISRLILISCLMDLKNFMDIOTCTMOLESSILCSPLPS 202  
 143 HDVTDNKLRLRSKNGKVLISRLILISCLMDLKNFMDIOTCTMOLESSILCSPLPS 191  
 203 LSLVSGYTKMDLVEFMLEDAPAVQVAGEGLTLPOFTLRDEKDLGCTKHYNTKFTCIEMK 262  
 192 ---SEGYTMNDLFEFMLSQGP-VQVAGEGLTLPOFTLRDEKDLGCTKHYNTKFTCIEMK 247  
 263 FHLERQMGYLLIOMTIPSLIYIISWVFWIMDAPARVIGITVTLMTTQSSGRAS 322  
 248 FHLERQMGYLLIOMTIPSLIYIISWVFWIMDAPARVIGITVTLMTTQSSGRAS 307  
 323 LPKVSYVAIDIMMAVCLLEFPAALLEYAAVNFVSRQKEFLRLRRORLEEDIIQ 382  
 308 LPKVSYVAIDIMMAVCLLEFPAALLEYAAVNFVSRQKEFLRLRRORLEEDVIRE 367  
 363 SREYRGYGLGHCLQARDGPMESGSIYS-PPAPAPLREGETTRKLYVD 431  
 368 SRFNFGYGMGHCLQVKGDTAVKATPANPLPQP---KDGDAIKKKFVD 413

Db

RESULT 3  
 S14816  
 glycine receptor alpha-2 chain variant A precursor - rat  
 N:Alternate names: neonatal glycine receptor  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 20-Aug-1999  
 C:Accession: S14816; S15776; JN0112  
 R:Akagi, H.; Hirai, K.; Hashinuma, F.  
 FEBS Lett. 281, 160-166, 1991  
 A:Title: Cloning of a glycine receptor subtype expressed in rat brain and spinal cord du  
 A:Reference number: S14816; MUID:91200276; PMID:1707830  
 A:Accession: S14816  
 A:Molecule type: DNA  
 A:Residues: 1-452 <AKA>  
 A:Cross-references: EMBL:X57281; NID:956743; PIDN:CAA40549.1; PID:956744  
 R:Kunze, J.; Kuryatov, A.; Maulet, Y.; Malosio, M.L.; Schmieden, V.; Betz, H.  
 FEBS Lett. 283, 73-77, 1991  
 A:Title: Alternative splicing generates two isoforms of the alpha-2 subunit of the inhib  
 A:Reference number: S15776; MUID:91243883; PMID:1645300  
 A:Accession: S15776  
 A:Status: not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 1-452 <KUH>  
 A:Cross-references: GB:X61159; NID:9286344; PIDN:CAA43471.1; PID:9286345

R:Kunze, J.; Schmieden, V.; Betz, H.  
 Neuron 5, 867-873, 1990  
 A:Title: A single amino acid exchange alters the pharmacology of neonatal rat glycine  
 A:Reference number: JN0112; MUID:91097798; PMID:2176511  
 A:Accession: JN0112  
 A:Molecule type: mRNA  
 A:Residues: 1-193, 'E', 195-452 <KU2>  
 C:Comment: Glycine reduces neuronal firing by activating this inhibitory glycine rece  
 C:Superfamily: acetylcholine receptor  
 C:Keywords: glycoprotein; ion channel; neurotransmitter receptor; transmembrane prote  
 F:1-27/Domain: signal sequence #status predicted <SIG>  
 F:28-452/Product: glycine receptor alpha-2 chain #status predicted <MAT>  
 F:253-280/Domain: transmembrane #status predicted <TM1>  
 F:287-304/Domain: transmembrane #status predicted <TM2>  
 F:316-342/Domain: transmembrane #status predicted <TM3>  
 F:424-441/Domain: transmembrane #status predicted <TM4>  
 F:72,103/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 74.5%; Score 1677; DB:2; Length 452;  
 Best Local Similarity 75.5%; Pred. No. 3.9e-136;  
 Matches 327; Conservative 32; Mismatches 48; Indels 26; Gaps 6;

1 MTLVPAATLSFLMLTLPGQVLLRYALAKE-EVKSQTRGSPSPDFLDKLMGRTSGYD 59  
 5 LVNITLALFAEFL-----GTNHFREAFCKDHDSRGKHPQSOTLSPSDFLDKLMGRTSGYD 59  
 60 ARIRNFKGPPVNTCNFINFSFSVTKTMDYRVNVEFLRQOMDPRLSREYDDSLD 119  
 60 ARIRNFKGPPVNTCNFINFSFSVTKTMDYRVNVEFLRQOMDPRLSREYDDSLD 119  
 120 DPSMDSIMKPDLEFANEKANFHEVTDNKLRLRFKNGNVLISRLILISCLMDLKNF 179  
 120 DPSMDSIMKPDLEFANEKANFHDVTDNKLRLRSKNGKVLISRLILISCLMDLKNF 179  
 180 PMDIOTCTMOLESSILCSPLPSLSLVSQYTKMDLVEFMLEDAPAVQVAGEGLTLPOFTLR 239  
 180 PMDIOTCTMOLESSILCSPLPSLSLVSQYTKMDLVEFMLEDAPAVQVAGEGLTLPOFTLR 224  
 180 PMDIOTCTMOLESSILCSPLPSLSLVSQYTKMDLVEFMLEDAPAVQVAGEGLTLPOFTLR 224  
 240 DEKDLGCTKHYNTKFTCIEMKFEHLERQMGYLLIOMTIPSLIYIISWVFWIMDAP 299  
 225 EEKELGYCTKHYNTKFTCIEMKFEHLERQMGYLLIOMTIPSLIYIISWVFWIMDAP 284  
 300 ARVIGITVTLMTTQSSGRASLPKVSYVAIDIMMAVCLLEFPAALLEYAAVNFVSRQ 359  
 285 ARVIGITVTLMTTQSSGRASLPKVSYVAIDIMMAVCLLEFPAALLEYAAVNFVSRQ 344  
 360 HKEFTLRRLRRORLEEDIIQESREYRGYGLGHCLQARDGPMESGSIYS-PPAPAPL 418  
 345 HKEFTLRRLRRORLEEDVIRESFNFGYGMGHCLQVKGDTAVKATPANPLPQP--- 401  
 419 LREGETTRKLYVD 431  
 402 -KDDADAIKKKFFVD 413

Db

RESULT 4  
 S18836  
 glycine receptor alpha-2 chain precursor variant B - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 19-Mar-1997 #sequence\_revision 23-Apr-1997 #text\_change 21-Aug-1998  
 C:Accession: S18836  
 R:Kunze, J.; Kuryatov, A.; Maulet, Y.; Malosio, M.L.; Schmieden, V.; Betz, H.  
 FEBS Lett. 283, 73-77, 1991  
 A:Title: Alternative splicing generates two isoforms of the alpha-2 subunit of the in  
 A:Reference number: S18836; MUID:91243883; PMID:1645300  
 A:Accession: S18836  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-452 <KUH>  
 C:Superfamily: acetylcholine receptor  
 C:Keywords: glycoprotein; membrane protein  
 F:1-27/Domain: signal sequence #status predicted <SIG>  
 F:28-452/Product: glycine receptor alpha-2 chain variant B #status predicted <MAT>

F:103/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 74.2%; Score 1671; DB 2; Length 452;

Best Local Similarity 75.1%; Pred. No. 1.3e-135; Mismatches 325; Conservative 33; Mismatches 49; Indels 26; Gaps 6;

```
OY 1 MTLVPAATSLFLLMTLPQVLLRVALAKE-EVKSCTGSGQSPSPDLKMGRTSGYD 59
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 5 LVNLTALFAFEL-----GTNHFREAFCKDHDNRSGKHPQSOTLSPSPDLKMGRTSGYD 59

OY 60 ARIRPNKGPVNVNTCNIFINSFSSVTKTMDYRVNVLFRQOMDPRLSREYDDSDL 119
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 60 ARIRPNKGPVNVNTCNIFINSFSGVETMDYRVNVLFRQOMDPRLSREYDDSDL 119
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

OY 120 DPSMLDSIMKPDLEFANEKANFHEVTTDNKLRIFKNGVLSIRLTLSCLMDLKNF 179
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 120 DPSMLDSIMKPDLEFANEKANFHDVTTDNKLRISKNGKLVLSIRLTLSCLMDLKNF 179

OY 180 PMDIQTCTMOLESSILCSPLPSLSVGYTKMDLVFEWLEDAVAOVAEGTLPOETLR 239
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 180 PMDIQTCTMOLE-----SFGYTMNDLIFEWLSDDGP-VQVAEGTLPOETLR 224

OY 240 DEKLGCTKHYNTGKFTCIEVKFHLEROMGYLIOMYIPSLLVILISWSFWINMDAAP 299
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 225 EEKELGCTKHYNTGKFTCIEVKFHLEROMGYLIOMYIPSLLVILISWSFWINMDAAP 284

OY 300 ARVGLGITVLTMTTQSSGRASLPKVSYYKAIDIMNAVCLLFYFALLLEYAALNFVSRO 359
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 285 ARVALGITVLTMTTQSSGRASLPKVSYYKAIDIMNAVCLLFYFALLLEYAALNFVSRO 344

OY 360 HKEFIRLRORRORLEEDIIQESRFYFRGYGLCHCLOARDGMEGSGIYS-PQPPAPL 418
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 345 HKEFIRLRORRORLEEDVRESRFNFSGYGMCHCLOKDGTAVKATPANPLPOPP--- 401

OY 419 LREGETTRKLYVD 431
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 402 -KDADAIRKKFVD 413
```

## RESULT 5

B49970

glycine receptor alpha-2 chain - mouse

C:Species: Mus musculus (house mouse)

C:Date: 17-Nov-1995 #sequence\_revision 17-Nov-1995 #text\_change 17-Mar-1999

C:Accession: B49970; S15802

R:Matzenbach, B.; Maulet, Y.; Sefton, L.; Courtier, B.; Avner, P.; Guenet, J.L.; Betz, H.

J. Biol. Chem. 269, 2607-2612, 1994

A:Title: Structural analysis of mouse glycine receptor alpha subunit genes. Identification

A:Reference number: A49970; MUID:94132024; PMID:7507926

A:Accession: B49970

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-451 <MAT>

A:Cross-references: GB:X75841

R:Kunze, J.; Kurtyakov, A.; Maulet, Y.; Malosio, M.L.; Schmieden, V.; Betz, H.

FEBS Lett. 283, 73-77, 1991

A:Title: Alternative splicing generates two isoforms of the alpha-2 subunit of the inhib

A:Reference number: S15776; MUID:91243883; PMID:1645300

A:Accession: S15802

A:Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 67-89 <KUH>

C:Genetics:

A:Gene: GIRA2

C:Superfamily: acetylcholine receptor

C:Keywords: neurotransmitter receptor; transmembrane protein

Query Match 73.6%; Score 1656.5; DB 2; Length 451;

Best Local Similarity 75.1%; Pred. No. 2.2e-134;

Mismatches 325; Conservative 33; Mismatches 48; Indels 27; Gaps 7;

```
OY 1 MTLVPAATSLFLLMTLPQVLLRVALAKE-EVKSCTGSGQSPSPDLKMGRTSGYD 59
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 LVNLTALFAFEL-----GTNHFREAFCKDHDNRSGKHPQSOTLSPSPDLKMGRTSGYD 59
```

```
Db 5 LVNLTALFAFEL-----GTNHFREAFCKDHDNRSGKHPQSOTLSPSPDLKMGRTSGY 58

OY 60 ARIRPNKGPVNVNTCNIFINSFSSVTKTMDYRVNVLFRQOMDPRLSREYDDSDL 119
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 59 ARIRPNKGPVNVNTCNIFINSFSGVETMDYRVNVLFRQOMDPRLSREYDDSDL 118

OY 120 DPSMLDSIMKPDLEFANEKANFHEVTTDNKLRIFKNGVLSIRLTLSCLMDLKNF 179
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 119 DPSMLDSIMKPDLEFANEKANFHDVTTDNKLRISKNGKLVLSIRLTLSCLMDLKNF 178

OY 180 PMDIQTCTMOLESSILCSPLPSLSVGYTKMDLVFEWLEDAVAOVAEGTLPOETLR 239
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 179 PMDIQTCTMOLE-----SFGYTMNDLIFEWLSDDGP-VQVAEGTLPOETLR 223

OY 240 DEKLGCTKHYNTGKFTCIEVKFHLEROMGYLIOMYIPSLLVILISWSFWINMDAAP 299
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 224 EEKELGCTKHYNTGKFTCIEVKFHLEROMGYLIOMYIPSLLVILISWSFWINMDAAP 283

OY 300 ARVGLGITVLTMTTQSSGRASLPKVSYYKAIDIMNAVCLLFYFALLLEYAALNFVSRO 359
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 284 ARVALGITVLTMTTQSSGRASLPKVSYYKAIDIMNAVCLLFYFALLLEYAALNFVSRO 343

OY 360 HKEFIRLRORRORLEEDIIQESRFYFRGYGLCHCLOARDGMEGSGIYS-PQPPAPL 418
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 344 HKEFIRLRORRORLEEDVRESRFNFSGYGMCHCLOKDGTAVKATPANPLPOPP--- 400

OY 419 LREGETTRKLYVD 431
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 401 -KDADAIRKKFVD 412
```

## RESULT 6

A23682

glycine receptor alpha-3 chain precursor - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 04-Oct-1991 #sequence\_revision 04-Oct-1991 #text\_change 20-Aug-1999

C:Accession: A23682

R:Kunze, J.; Schmieden, V.; Betz, H.

J. Biol. Chem. 265, 22317-22320, 1990

A:Title: Identification and functional expression of a novel ligand binding subunit o

A:Reference number: A23682; MUID:91093073; PMID:2176214

A:Accession: A23682

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-464 <KUH>

A:Cross-references: GB:M5250; GB:M38385; NID:q204882; PIDN:AAA63492.1; PID:q204883

A:Superfamily: acetylcholine receptor

C:Keywords: transmembrane protein

Query Match 70.7%; Score 1592; DB 2; Length 464;

Best Local Similarity 71.3%; Pred. No. 8e-129;

Mismatches 313; Conservative 34; Mismatches 50; Indels 42; Gaps 7;

```
OY 9 LSTLLMTLPQVLLRVALAKEEVKSGTKSGQSPSPDLKMGRTSGYARIRPNFKG 68
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 11 LSGYFWE-----AALLSLVATKETSARSASPMSPDLKMGRTSGYARIRPNFKG 67

OY 69 PPVAVTNCIFINSFSSVTKTMDYRVNVLFRQOMDPRLSREYDDSDLDPMSLSIW 128
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 68 PPVAVTNCIFINSFSGVETMDYRVNVLFRQOMDPRLSREYDDSDLDPMSLSIW 127

OY 129 KPDLFEANEKANFHEVTTDNKLRIFKNGVLSIRLTLSCLMDLKNFPMDIQTCTM 188
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 128 KPDLFEANEKANFHEVTTDNKLRIFKNGVLSIRLTLSCLMDLKNFPMDIQTCTM 187

OY 189 QLESSSILCSPLPSLSVGYTKMDLVFEWLEDAVAOVAEGTLPOETLRDEKDLGCT 248
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 188 QLE-----SFGYTMNDLIFEWLSDDGP-VQVAEGTLPOETLRDEKDLRVCY 232

OY 249 KHYNTGKFTCIEVKFHLEROMGYLIOMYIPSLLVILISWSFWINMDAAPARVGLGTT 308
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 233 KHYNTGKFTCIEVKFHLEROMGYLIOMYIPSLLVILISWSFWINMDAAPARVGLGTT 292
```







```

F:331-354/Domain: transmembrane #status predicted <TM3>
F:478-495/Domain: transmembrane #status predicted <TM4>
F:54_242_387/binding site: carboxylate (Asp) (covalent) #status predicted
F:383-197_243-255/disulfide bonds: #status predicted
F:365_410/binding site: phosphate (Thr) (covalent) #status predicted
F:435/binding site: phosphate (Tyr) (covalent) #status predicted

```

Query Match	38.3%	Score 863.5;	DB 2;	Length 496;
Best Local Similarity	47.8%;	Pred. No. 3.5e-66;		
Matches 187;	Conservative 63;	Mismatches 88;	Indels 53;	Gaps 11

```

QY 9 L$FLLMTLPGQVLLALVAALAKEEVSQGTGSQ-----PM$D$DL 50
D$ 7 V$F$ILL$M$-----L$F$D$A$C$K$E$S$S$K$G$K$G$K$Q$Y$C$Q$S$Q$A$E$D$A$R$V$P$N$S$T$N$LN$R 62
QY 51 LM$RT$G$Y$D$A$R$IN$P$N$K$G$P$P$V$N$T$C$N$IF$N$F$S$V$T$K$T$M$D$Y$R$V$N$V$E$LR$Q$M$D$P$R$L$--- 107
D$ 63 LL$---V$Y$D$P$R$IN$P$N$K$G$I$P$D$V$Y$V$N$IF$N$S$G$T$Q$E$T$M$D$Y$R$V$N$IF$R$Q$M$D$P$R$LL$K$P 119
QY 108 -S$T$R$E$Y$P$D$S$LD$D$P$M$LD$S$IK$P$D$L$F$E$A$N$K$G$A$N$F$E$V$T$T$D$N$K$L$R$IF$E$K$G$N$V$Y$S$IR$L 166
D$ 120 S$D$R$---G$S$D$A$L$T$V$D$P$M$Y$K$C$LM$K$P$D$L$F$E$A$N$K$S$A$N$F$H$D$V$Q$E$N$ILL$F$R$G$D$V$Y$S$M$R$L 177
QY 167 T$LL$C$L$M$D$LN$K$N$P$M$H$I$Q$T$C$T$M$Q$E$S$S$T$L$C$P$L$S$L$S$Y$G$Y$M$K$D$V$E$F$M$E$D$A$P$Q 226
D$ 178 S$T$L$C$P$D$LD$T$F$P$M$O$T$O$R$C$K$M$Q$E$-----S$F$Y$T$D$D$R$F$T$M$O$S$D$P$-V$Q 222
QY 227 V$A$G$L$T$P$O$F$-I$R$D$E$D$G$C$C$T$K$N$-N$T$G$R$F$C$T$E$Y$E$F$H$E$R$O$M$G$Y$LL$M$Y$P$S$LL$Y 284
D$ 223 L-$E$K$A$LP$Q$D$IK$E$D$E$Y$G$N$C$T$K$Y$K$G$Y$T$Y$C$V$E$F$T$L$R$O$G$Y$F$M$Q$V$N$P$T$LL$Y 281
QY 285 I$L$S$W$F$W$IN$M$D$A$P$A$R$V$G$I$T$V$L$M$T$O$S$G$S$R$A$L$P$K$V$S$Y$V$K$A$ID$IM$A$V$C$LL$E$V$F 344
D$ 282 V$L$S$M$L$F$W$IN$P$D$A$S$A$R$V$P$G$I$F$S$V$L$A$S$E$C$T$T$L$A$E$L$P$K$V$S$Y$V$K$A$D$V$W$L$I$A$C$LL$E$G$F 341
QY 345 A$L$E$Y$A$T$-----N$F$S$R$O$K$E$F$IL$R$R$O$R 371
D$ 342 A$S$L$E$T$A$V$O$V$M$LN$N$P$K$R$E$A$K$A$R$I$A$A$Q 372

```

RESULT 12  
G02031  
glycine receptor beta subunit - human  
C:Species: Homo sapiens (man)  
C:Date: 21-Dec-1996 #sequence\_revision 06-Jun-1997 #text\_change 20-Aug-1999  
C:Accession: G02031  
R:Handford, C.A.; Schofield, P.R.  
submitted to the EMBL Data Library, August 1995  
A:Reference number: G03080  
A:Accession: G02031  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1..497 <H>  
A:Cross-references: EMBL:U03267; NID:g992686; PID:AAB3750.1; PID:g992687  
C:Genetics:  
A:Gene: GDB:GLRB  
A:Cross-references: GDB:5875373; OMT:138492  
A:Map position: 4q32-4q32  
A:Superfamily: acetylcholine receptor

```
Query Match      38.0%; Score 85.5; DB 2; Length 497;  
Best Local Similarity 47.4%; Pred. NO. 1.7e-65;  
Matches 186; Conservative 64; Mismatches 89; Indels 53; Gaps 11  
  
QY      8 TLSELMTTLPGLGVALRALAKEEKSCTKSGSPM-----SPSDFLD 49  
| : | : | : | : | : | : | : | : | : | : | : | : | : |  
Db      6 TTTALLILISL----WVEEAYSKEKSSKKGGKKKKOYLCPSSOASADLARVPANSTNLIN 61  
| : | : | : | : | : | : | : | : | : | : | : | : | : |  
  
QY      50 KIMRTSGDYARIRPNKGPPVNYTCNIFINSFSVTPTTNDYRVNVLROOMNDPRL-- 107  
| : | : | : | : | : | : | : | : | : | : | : | : | : |  
Db      62 RLL---VSYPDIRIRPNKGIPVDVVYNFINFSFGSIQETMDRYRNITFLRCKMKNDPRKL 118  
| : | : | : | : | : | : | : | : | : | : | : | : | : |  
  
QY      108 --SYREYPDSDLDPSMLDSIMKPDLFEANEKGANFHEVTTDNKLRIFKNGNVLYSIR 165
```

Db 119 pSOFr--GSDALVDPPTMYKCKMRKDDLEFAENKSNFNFIDYQEBILLFIRDDGVLSMR 176

Qy 166 LTIILISGLMDKNFPMDIOTCTMQLSESSILCSPILPSLISVGYTMDVLFEMLEDA PAY 225

Db 177 LSTLSLCPDLTLEFPMDIOCKRMQLE-----SEGTTDDLRFTMQSGDP-V 221

Qy 226 QVAEGITLFOF-ILRDEKDLGCCTAHY-NTGCTFCIEYKPHLEHOMQVYLIOAYISLIL 283

Db 222 QL-EKLTALPOFDLKREDIEYNGCTKYKYGCTGYTCVEIFPLRQOVFFYMGVAPATLLI 280

Qy 284 VILSWSPFNMDAAARVGLGITTVLMMTQSSGRSLDEPKSVYVAIDIMAAVCLFFV 343

Db 281 VVLSWLSFMINPDASAARVPLGISVLSLASCSCTTLAAELPKSVYVALADVMVLIACLIFG 340

Qy 344 FAALLFEYAAI-----NFVSQKHKEFIRLRRROR 371

Db 341 FASLVEIAAVQVMLNPNKRFVEAEKRIKAAID 372

RESULT 13

gamma-aminobutyric acid/benzodiazepine receptor A beta chain precursor - great pond snail

S:17785

N:Alternate names: GABA(A) receptor

C:Species: *Lymanea stagnalis* (great pond snail)

C:Date: 16-Sep-1992 #sequence\_revision 16-Sep-1992 #text\_change 20-Aug-1999

C:Accession: S17785; A60170

R:Harvey, R.J.; Vreugdenhil, E.; Zaman, S.H.; Bhandal, N.S.; Usherwood, P.N.R.; Barnard

EMBO J. 10, 3219-3245, 1991

A:Title: Sequence of a functional invertebrate GABA(A) receptor subunit which can form

A:Reference number: S17785; MUID:92007774; PMID:1655414

A:Accession: S17785

A:Status: nucleic acid sequence not shown; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-499 <HAR1>

A:Cross-references: GB:X58638; NID:g9626

A:Note: the translated sequence in GenBank entry LSGABA, release 111.0, (PID:g9627)

R:Harvey, R.J.; Vreugdenhil, E.; Barnard, E.A.; Darlison, M.G.

Biochem. Soc. Trans. 18, 438-439, 1990

A:Title: Cloning of genomic and cDNA sequences encoding an invertebrate gamma-aminobu

A:Reference number: A60170; MUID:90323312; PMID:2164988

A:Accession: A60170

A:Status: nucleic acid sequence not shown; not compared with conceptual translation

A:Molecule type: DNA; mRNA

A:Residues: 201-231, 'L', 233-334 <HAR2>

C:Superfamily: acetylcholine receptor

C:Keywords: glycoprotein; neurotransmitter receptor; transmembrane protein

F:1-23/Domain: signal sequence #status predicted <SIG>

F:24-499/Product: gamma-aminobutyric acid/benzodiazepine receptor A beta chain #status predi

F:32, 98, 106, 152, 334, 367, 489/Binding site: carbonyl(ate) (asn) (covalent) #status predi

[illegible]

Db 257 YLPSILVMSVSWFMINHEATSAVALGITTTLMTTISNGVNSSLPRISYVKAIDYL 316  
QY 337 AVCLLFVFAALLEVAALNFV---SROHKEFIRLRPROR--RORLEE 377  
Db 317 VMCEVFVFAALLEVAALNFV---SROHKEFIRLRPROR--RORLEE 377  
RESULT 14  
JH0359  
gamma-aminobutyric acid receptor A beta-4' chain precursor - chicken  
N:Contains: gamma-aminobutyric acid receptor A beta-4 chain, long splice form; gamma-am  
C:Species: Gallus gallus (chicken)  
C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 20-Aug-1999  
C:Accession: JH0359; JH0360; B61282; A61282; S29694; S29696; S29695  
R:Bateson, A.N.; Lasham, A.; Darlison, M.G.  
J. Neurochem. 56, 1437-1440, 1991  
A:Title: Gamma-aminobutyric acid A receptor heterogeneity is increased by alternative sp  
A:Reference number: JH0359; MUID:91162222; PMID:1848280  
A:Accession: JH0359  
A:Molecule type: mRNA  
A:Residues: 1-488 <BAT>  
A:Cross-references: EMBL:X56647  
A:Experimental source: brain  
A:Accession: JH0360  
A:Molecule type: mRNA  
A:Residues: 1-360,365-488 <BA2>  
A:Cross-references: EMBL:X56648  
A:Experimental source: brain  
R:Lasham, A.; Bateson, A.N.; Darlison, M.G.  
Biochem. Soc. Trans. 19, 9S, 1991  
A:Title: Alternative splicing increases GABA-A receptor heterogeneity.  
A:Reference number: A61282; MUID:91244095; PMID:1645322  
A:Accession: B61282  
A:Molecule type: mRNA  
A:Residues: 357-366 <LAS>  
A:Accession: A61282  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 357-360,365-366 <LA2>  
R:Bateson, A.N.; Lasham, A.; Darlison, M.G.  
submitted to the EMBL Data Library, November 1990  
A:Description: Gamma-aminobutyric acid A receptor heterogeneity is increased by alternat  
A:Reference number: S29694  
A:Accession: S29694  
A:Molecule type: DNA  
A:Residues: 364-488 <BAW>  
A:Cross-references: EMBL:X56646; NID:g62912; PIDN:CAA39968.1; PID:g62913  
A:Accession: S29696  
A:Molecule type: mRNA  
A:Residues: 1-405 <BA3>  
A:Cross-references: EMBL:X56648  
A:Comment: The gamma-aminobutyric acid A receptor is the major inhibitory neurotransmit  
C:Superfamily: acetylcholine receptor  
C:Keywords: alternative splicing; glycoprotein; neurotransmitter receptor; phosphoprotei  
F:1-25/Domain: signal sequence #status predicted <SIG>  
F:26-488/Product: gamma-aminobutyric acid A receptor beta-4' chain #status predicted <M  
F:26-360,365-488/Product: gamma-aminobutyric acid A receptor beta-4 chain #status predic  
F:24-26/Domain: transmembrane #status predicted <TM1>  
F:271-292/Domain: transmembrane #status predicted <TM2>  
F:304-336/Domain: transmembrane #status predicted <TM3>  
F:466-487/Domain: transmembrane #status predicted <TM4>  
F:32,104,173,195/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:160-174/Disulfide bonds: #status predicted  
F:448/Binding site: phosphate (Ser) (covalent) #status predicted  
Query Match 32.2%; Score 726; DB 2; Length 488;  
Best Local Similarity 38.0%; Pred. No. 2,2e-54;  
Matches 156; Conservative 79; Mismatches 124; Indels 52; Gaps 10;

QY 43 SPS-----DFLDKLMGRTSGYDARIRPNFKGPVNVNTGIFINFSSTVTKTMDRV 94  
Db 27 SPSGNGISVNAKEIYDKILL---KGYDVRRLRPFQGNPVTGSHISSIDQSEVNDYTI 83  
QY 95 NVFLRQONDRPLSTREYRPPDSDLDPSMLDSINKPDLFFANEGANFHEVTTDNKLLRI 154  
Db 84 TMYFOOSWRDKRLAYNDLP-LNLTLDNRVADQMLPDTYFPLNDKSPFHGVTVANRMIRL 142  
QY 155 FKNGNVLYSIRLTLLSLCMLDKNPPMDIQCTMOLESSILCSPLPSLSVGYTMKD 214  
Db 143 HPDGTVLGLGRTTAAACMDLRRPDLDOQCTLEIE-----SYGVYDDI 188  
QY 215 VFEMLDAPAVQVAEGTLTPQFILLDEKDLCCCTKHYNTGKFTCEVFKHLEROMGYLI 274  
Db 189 VFFQNGNSAVTGNHEVLELPQFTIIQR-LVSREVFTTSGYLRLSLFRKRNIGVITL 247  
QY 275 QMYTSLIIVLTVSWFMINNDAPARVGLGITTIVLTMTTSSGSRASLPVSYVKAIDI 334  
Db 248 QYMPSTILITLTVSWFMINNDASARVALGVTTVLTMTTINTHLRFLPKIPYVKAIDV 307  
QY 335 WMAVCLLFVFAALLEVAALNFV-----SROHKEFIRL-----RRQRORLEEDIOE 382  
Db 308 YLMGCFVVFALALEYAVNTIIFGRCRPOOKOSERISKANNEHRERKEKRVQDVPY 367  
QY 383 SRFYFRGYGLCHCLQARD-----GG-----PMESGIGYSPQAPAPLUREG 422  
Db 368 GNILLSTLDMNNELLATDMSSVGSRSVMSFESGIGQFRRPLAS--RDG 416

## RESULT 15

S33744  
cyclo diene insecticide resistance protein - yellow fever mosquito  
C:Species: Aedes aegypti (yellow fever mosquito)  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 20-Aug-1999  
C:Accession: S33744  
R:Thompson, M.; Sholtski, F.; french-Constant, R.  
FEBS Lett. 325, 187-190, 1993  
A:Title: Cloning and sequencing of the cyclo diene insecticide resistance gene from th  
A:Reference number: S33744; MUID:93307492; PMID:8391473  
A:Accession: S33744  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-533 <THO>  
A:Cross-references: GB:U28803; NID:g881589; PIDN:AAA68961.1; PID:g881590  
C:Superfamily: acetylcholine receptor  
Query Match 32.1%; Score 724; DB 2; Length 533;  
Best Local Similarity 38.6%; Pred. No. 3.7e-54;  
Matches 146; Conservative 80; Mismatches 122; Indels 30; Gaps 6;

QY 32 VKSGTKSGQSPSPDFDLKMGRTSGYDARIRPNFKGPVNVNTGIFINFSSTVTKTMD 91  
Db 39 VLAGTGGSMGLDGVNISAIDLDSFVGDKRRVPPGVEGVYTMVLSISSVSEVLMD 98  
QY 92 YRVNVFLRQONDRPLSTREYRPPD-DSLDLPSMLDSINKPDLFFANEGANFHEVTTDNK 150  
Db 99 FTLDYFQFOFTDRLAARAKRGVETLSVSGSEFININVPDTFFVNEQSFHIAITTSNE 158  
QY 151 LLRIFFKNGNVLYSIRLTLLSLCMLDKNPPMDIQCTMOLESSILCSPLPSLSVGYT 210  
Db 159 FIRVHSSGSTRIRLTITTASCMPGLQYFPMRDLCHIEIE-----SFGYT 204  
QY 211 MKDLVFEVLEDPANQVAEGTLTPQFILLDEKDLCCCTKHYNTGKFTCEVFKHLEROMG 270  
Db 205 MDRIRYFVKDGLSVGMSSEVELLPQFVLGHRQ-RATEINLTGVSRLACIOPFRVRSG 263  
QY 271 YLLQMYTSLIIVLTVSWFMINNDAPARVGLGITTIVLTMTTSSGSRASLPVSYVYK 330  
Db 264 YLLDIYIIPSGIIVTVSWFMINNDAPARVGLGITTIVLTMTTSSGSRASLPVSYVYK 323  
QY 331 AIDIMAVCLLFVFAALLEVAALNFVSRQ---HKEFIRLR--RORRORLEE----- 377

Tue Jul 1 10:05:51 2003

us-10-075-846-4.rpr

Page 8

Db 324 SIDYLGICFVWVFSLSEIYATVGIMAKNIQIGKRFMAIQIAEOKKOADANHPPPP 363

Qy 378 -DIIQESRFYERGYGLGH 394

Db 384 PPVSDSHSHGHGSHGH 401

Search completed: June 25, 2003, 17:17:34  
Job time : 29.9262 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 25, 2003, 17:12:59 ; Search time 17.3557 Seconds

(without alignments)  
1029.995 Million cell updates/sec

Title: US-10-075-846-4

Perfect score: 2252

Sequence: 1 MTTLVPAFLSFLMLTLPQ.....PPAPAPLREGETTRKLYVD 431

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1682.5	74.7	452	1 GRA2_HUMAN	P23416 homo sapien
2	1678	74.5	337	1 GRA4_MOUSE	Q61603 mus musculu
3	1670	74.2	452	1 GRA2_RAT	P22771 ratu
4	1602.5	71.2	465	1 GRA3_HUMAN	O75311 homo sapien
5	1592	70.7	464	1 GRA3_RAT	P24524 ratu
6	1580.5	70.2	449	1 GRA1_HUMAN	P23415 homo sapien
7	1580	69.8	457	1 GRA1_BOVIN	P57695 bos taurus
8	1573	69.8	444	1 GRA1_BRARE	O93430 brachydanc
9	1569.5	69.7	457	1 GRA1_MOUSE	Q64018 mus musculu
10	1568.5	69.6	457	1 GRA1_RAT	P07727 ratu
11	871	38.7	496	1 GRB_MOUSE	P48168 mus musculu
12	863.5	38.3	496	1 GRB_RAT	P20781 ratu
13	855.5	38.0	497	1 GRB_HUMAN	P48167 homo sapien
14	742	32.9	499	1 GAB_LYMT	P26714 lymnaea sta
15	740	32.2	488	1 GAB4_CHICK	P24045 gallu
16	720	32.0	466	1 GAB3_DROME	P25123 drosophila
17	715	31.7	496	1 GAB2_DROME	O08832 drosophila
18	712	31.6	474	1 GAB2_HUMAN	P47870 homo sapien
19	712	31.6	474	1 GAB2_MOUSE	P15442 mus musculu
20	697	31.0	474	1 GAB1_MOUSE	P50571 mus musculu
21	696	30.9	473	1 GAB3_HUMAN	P28472 homo sapien
22	695.5	30.9	474	1 GAB1_RAT	P15431 ratu
23	694.5	30.8	474	1 GAB1_BOVIN	P08220 bos taurus
24	692	30.7	473	1 GAB3_MOUSE	P15435 mus musculu
25	691.5	30.7	474	1 GAB3_HUMAN	P18505 homo sapien
26	687.5	30.5	476	1 GAB3_CHICK	P19019 gallu
27	676.5	29.7	452	1 GAD_HUMAN	O14764 homo sapien
28	669	29.0	638	1 GAAT_MOUSE	O91111 mus musculu
29	664.5	29.5	449	1 GAAT_RAT	P18506 ratu
30	663	29.4	440	1 GAAT_MOUSE	O09028 ratu
31	653.5	29.0	449	1 GAD_MOUSE	P22933 mus musculu
32	652.5	29.0	440	1 GAAT_HUMAN	O00531 homo sapien
33	644.5	28.6	465	1 GAR2_HUMAN	P28476 homo sapien

34	640.5	28.4	453	1 GA46_HUMAN	Q16445 homo sapien
35	640	28.4	474	1 GAR1_MOUSE	P56475 mus musculu
36	638.5	28.4	453	1 GAR6_MOUSE	P16305 mus musculu
37	637.5	28.3	632	1 GAAT_HUMAN	O90888 homo sapien
38	636	28.2	474	1 GAR1_RAT	P50572 ratu
39	634.5	28.2	453	1 GA6_RAT	P30191 ratu
40	634	28.2	474	1 GAC2_CHICK	P21548 gallu
41	633.5	28.1	465	1 GA6_CHICK	O90845 gallu
42	631.5	28.0	473	1 GAR1_HUMAN	P24046 homo sapien
43	631.5	28.0	552	1 GA44_MOUSE	O96654 mus musculu
44	631	28.0	552	1 GA44_RAT	P28471 ratu
45	629	27.9	467	1 GAC3_MOUSE	P27681 mus musculu

## ALIGNMENTS

RESULT 1  
ID GRA2\_HUMAN STANDARD; PRT; 452 AA.  
AC P23416; 099862;  
DT 01-NOV-1991 (Rel. 20, Created)  
DT 01-NOV-1991 (Rel. 20, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Glycine receptor alpha-2 chain precursor.  
GN GLRA2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90183975; PubMed=2155780;  
RA Greeningloh G., Schmieden V., Schofield P.R., Seebury P.H.,  
RA Siddique T., Khandas T.K., Becker C.M., Betz H.;  
RT Alpha subunit variants of the human glycine receptor: primary  
RT structures, functional expression and chromosomal localization of the  
RT corresponding genes.";  
RT EMBO J. 9:771-776(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=9837669; PubMed=9674912;  
RA Cummings C.J., Dahle E.J.R., Zoghbi H.Y.;  
RA "Analysis of the genomic structure of the human glycine receptor  
RT alpha-2 subunit gene and exclusion of this gene as a candidate for  
RL Rett syndrome.";  
RL Am. J. Med. Genet. 78:176-178(1998).  
RN [3]  
RP SEQUENCE OF 1-360 FROM N.A.  
RX MEDLINE=97129407; PubMed=8973915;  
RA Monani U.R., Burghes A.H.M.;  
RT "Structure of the human alpha 2 subunit gene of the glycine receptor:  
RT use of vectorette and Alu-exon PCR.";  
RL Genome Res. 6:1200-1206(1996).  
CC - FUNCTION: THE GLYCINE RECEPTOR IS A NEUROTRANSMITTER-GATED ION  
CC CHANNEL. BINDING OF GLYCINE TO ITS RECEPTOR INCREASES THE CHLORIDE  
CC CONDUCTANCE AND THUS PRODUCES HYPERPOLARIZATION (INHIBITION OF  
CC NEURONAL FIRING)  
CC - SUBUNIT: PENTAMER COMPOSED OF ALPHA (48 kDa) AND BETA (58 kDa)  
CC SUBUNITS. A THIRD TYPE OF SUBUNIT (93 kDa) IS A PERIPHERAL  
CC MEMBRANE PROTEIN ASSOCIATED WITH THE CYTOPLASMIC DOMAINS OF THE  
CC RECEPTOR CORE.  
CC - SUBCELLULAR LOCATION: Integral membrane protein.  
CC - ALTERNATIVE PRODUCTS: 2 ISOFORMS: ALPHA-2\* (SHOWN HERE) AND ALPHA-  
CC 2B; ARE PRODUCED BY ALTERNATIVE SPLICING.  
CC - SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.  
CC - THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: X52008: CA36257.1: -

DR EMBL: AF053495: AAC35290.1: JOINED.

DR EMBL: AF053487: AAC35290.1: JOINED.

DR EMBL: AF053488: AAC35290.1: JOINED.

DR EMBL: AF053489: AAC35290.1: JOINED.

DR EMBL: AF053490: AAC35290.1: JOINED.

DR EMBL: AF053491: AAC35290.1: JOINED.

DR EMBL: AF053492: AAC35290.1: JOINED.

DR EMBL: AF053493: AAC35290.1: JOINED.

DR EMBL: AF053494: AAC35290.1: JOINED.

DR EMBL: U77731: AAB38272.1: JOINED.

DR EMBL: U77724: AAB38272.1: JOINED.

DR EMBL: U77725: AAB38272.1: JOINED.

DR EMBL: U77726: AAB38272.1: JOINED.

DR EMBL: U77727: AAB38272.1: JOINED.

DR EMBL: U77728: AAB38272.1: JOINED.

DR EMBL: U77729: AAB38272.1: JOINED.

DR EMBL: U77730: AAB38272.1: JOINED.

DR EMBL: U77731: AAB38273.1: JOINED.

DR EMBL: U77725: AAB38273.1: JOINED.

DR EMBL: U77726: AAB38273.1: JOINED.

DR EMBL: U77727: AAB38273.1: JOINED.

DR EMBL: U77728: AAB38273.1: JOINED.

DR EMBL: U77729: AAB38273.1: JOINED.

DR EMBL: U77730: AAB38273.1: JOINED.

DR PIR: S12381: S12381.

DR GeneW: HGNC:4327: GLRA2.

DR MIM: 305990: -

DR InterPro: IPR000188: GABA\_A\_receptor.

DR InterPro: IPR001175: Neur\_channel.

DR Pfam: PF02931: Neur\_chan\_LBD. 1.

DR Pfam: PF02932: Neur\_chan\_memb. 1.

DR PRINTS: PR00252: NRIONCHANNEL.

DR TIGRFAMs: TIGR00860: LIC. 1.

DR PROSITE: PS00236: NEUROT\_R\_ION\_CHANNEL. 1.

DR KX Receptor: Postsynaptic membrane; Ionic channel; Glycoprotein; Signal; Transmembrane; Multigene family; Alternative splicing.

FT SIGNAL 1 27

FT CHAIN 28 452

FT DOMAIN 28 253

FT TRANSMEM 254 280

FT TRANSMEM 287 304

FT TRANSMEM 319 342

FT DOMAIN 343 423

FT TRANSMEM 424 441

FT DISULFID 172 186

FT CARBOHYD 72 72

FT CARBOHYD 103 103

FT VARSPLIC 85 86

SEQUENCE 452 AA: 52002 MW: 95E74B3F9A555EA CAC64;

Query Match 74.7%; Score 1682.5; DB 1; Length 452;

Best Local Similarity 79.0%; Pred. No. 8e-136;

Matches 324; Conservative 29; Mismatches 36; Indels 21; Gaps 5;

OY 24 RVALAKE-EVKSGRKSGSPSPDLKMGRTSGYARIRPNKGPVNTVCNIFNSF 82

DB 23 RTACCKDHSRSGKOPSTLSPDLKMGRTSGYARIRPNKGPVNTVCNIFNSF 82

OY 83 SSVTKTMDYRVNLFQNMNDPLSYREYPPDSLDLPSMLDSIMKPDLEFFANEKGNF 142

DB 83 GSVEETMDYRVNLFQNMNDPLSYREYPPDSLDLPSMLDSIMKPDLEFFANEKGNF 142

OY 143 HEVTTDNKLLRIFKNGVLYSIRLTLLISCLMDKNFPMQIOTCTMOLESIIICSPUPS 202

DB 143 HDVTTDNKLLRISKNGVLYSIRLTLLISCLMDKNFPMQIOTCTMOLE----- 191

OY 203 LSLSVGTMDLVFEWLEDAVAOVASGLTLPQILRDEKDLGGCTKHYNFGKFTCLEVK 262

DB 192 ---SFGYTMNDLIFEMLSDDP-VQVABSLTLPQILRDEKDLGGCTKHYNFGKFTCLEVK 247

OY 263 FHLEKMGYLYLQMTPELIVIIISWSFWINMDAARVGLGTTVLTMTQSSGSRAS 322

DB 248 FHLEKMGYLYLQMTPELIVIIISWSFWINMDAARVGLGTTVLTMTQSSGSRAS 307

OY 323 LPKSVYKAIDIMAVCLFEVFAALLEYAALNFVSRQKKEFRLRRQRORLEEDIIQ 382

DB 308 LPKSVYKAIDIMAVCLFEVFAALLEYAALNFVSRQKKEFRLRRQRORLEEDIIQ 367

OY 383 SRFYFRGYLCHCLQARDGPEMGSGIYS-PQPAFLREGETTRKLYVD 431

DB 368 SRFNPSGYMGHCLQVKGDAVKAATPAAPLQPP-----KDDDAIKKFFVD 413

RESULT 2

GRA4\_MOUSE

ID GRA4\_MOUSE STANDARD: PRT: 337 AA.

AC 061603:

DT 15-DEC-1998 (Rel. 37, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Glycine receptor alpha-4 chain precursor (Fragment).

GN GLRA4.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_Taxid=10090;

RN (1)

RP SEQUENCE FROM N.A.

RC STRAIN=BALB/c;

RA MEDLINE=94132024; PubMed=7507926;

RA Matzebach B., Maulet Y., Seflon L., Courtier B., Avner P.,

RA Guenet J.-L., Betz H.,

RT "Structural analysis of mouse glycine receptor alpha subunit genes. Identification and chromosomal localization of a novel variant.";

RL J. Biol. Chem. 269:2607-2612(1994).

CC -1- FUNCTION: THE GLYCINE RECEPTOR IS A NEUROTANSITTER-GATED ION CHANNEL. BINDING OF GLYCINE TO ITS RECEPTOR INCREASES THE CHLORIDE CONDUCTANCE AND THIS PRODUCES HYPERPOLARIZATION (INHIBITION OF NEURONAL FIRING).

CC -1- SUBUNIT: PENTAMER COMPOSED OF ALPHA (48 kDa) AND BETA (58 kDa) SUBUNITS. A THIRD TYPE OF SUBUNIT (93 kDa) IS A PERIPHERAL MEMBRANE PROTEIN ASSOCIATED WITH THE CYTOPLASMIC DOMAINS OF THE RECEPTOR CORE.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.

CC -1- MISCELLANEOUS: THE ALPHA SUBUNIT BINDS STYCHNINE.

CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: X75850: CA53468.1: -

DR EMBL: X75851: CA53468.1: JOINED.

DR EMBL: X75852: CA53468.1: JOINED.

DR EMBL: X75853: CA53468.1: JOINED.

DR MGD: MGI:95750: Glra4.

DR InterPro: IPR001188: GABA\_A\_receptor.

DR InterPro: IPR001175: Neur\_channel.

DR Pfam: PF02931: Neur\_chan\_LBD. 1.

DR Pfam: PF02932: Neur\_chan\_memb. 1.

DR TIGRFAMs: TIGR00860: LIC. 1.

DR PROSITE: PS00236: NEUROT\_R\_ION\_CHANNEL. 1.

DR Receptor: Postsynaptic membrane; Ionic channel; Glycoprotein; Signal; Transmembrane; Multigene family.

FT NON\_TER 1 10

FT SIGNAL <1 10

POTENTIAL.

```

FT CHAIN 11 >337 GLYCINE RECEPTOR ALPHA-4 CHAIN.
FT DOMAIN 11 234 EXTRACELLULAR (PROBABLE).
FT TRANSMEM 235 255 POTENTIAL.
FT TRANSMEM 294 314 POTENTIAL.
FT DISULFID 148 162 BY SIMILARITY.
FT DISULFID 209 220 BY SIMILARITY.
FT CARBOHYD 48 48 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT NON_TER 337 337
SQ SEQUENCE 337 AA; 38774 MW; 9A3EDD7230578582 CRC64;

Query Match 74.58; Score 1678; DB 1; Length 337;
Best Local Similarity 92.08; Pred. No. 1.3e-135;
Matches 323; Conservative 7; Mismatches 7; Indels 14; Gaps 1;

OY 25 VALAKEEKGSGTSGSPSPDFDLKMGRTSGYDARIRPKGPPVNVTCNIFINSFSS 84
DB 1 VALAKEEKGSGKSGSPSPDFDLKMGRTSGYDARIRPKGPPVNVTCNIFINSFSS 60
OY 85 VTKTMDYRVNVFLRQOMNDPRLSYREXPDDSLDLPSMDSIMKPDLEFANEKANFHE 144
DB 61 VTEETMDYRVNVFLRQOMNDPRLAYREXPDDSLDLPSMDSIMKPDLEFANEKANFHE 120
OY 145 VTTDNKLRIRFKKNVYSTRILTLISCLNDLKNFPHDIOTCTMOLESSSLCSPLESL 204
DB 121 VTTDNKLRIRFKKNVYSTRILTLISCLNDLKNFPHDIOTCTMOLE----- 167
OY 205 LSVGYTKKDLFEWLEDAVAQVAEGTLTPOFILRDEKLGCCRKHNTGKTCIEYKFH 264
DB 168 -SEGYTMNDLMEFLEDAVAQVAEGTLTPOFILRDEKLGCCRKHNTGKTCIEYKFH 226
OY 265 LEROMGYLLIOMYIPSLIIVLISVSWFWINMDAARVGLGITTVLMTTQSSGSRASLP 324
DB 227 LEROMGYLLIOMYIPSLIIVLISVSWFWINMDAARVGLGITTVLMTTQSSGSRASLP 286
OY 325 KVSIVKADIDIMAVCLLFFPALLEVAALNFVSRHKEFTILRRORRORL 375
DB 287 KVSIVKADIDIMAVCLLFFPALLEVAALNFVSRHKEFTILRRORRORRORL 337

RESULT 3
GRA2_RAT STANDARD: PRT; 452 AA.
ID GRA2_RAT STANDARD: PRT; 452 AA.
AC P22771; 091w28;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Glycine receptor alpha-2* chain precursor (Neonatal isoform)
DE (Glycine receptor strychnine binding subunit).
GN GLRA2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND MUTAGENESIS OF GLU-194.
RC TISSUE-Brain;
RX MEDLINE=91097798; PubMed=2176511;
RA Kuhse J., Schmieden V., Betz H.;
RT "A single amino acid exchange alters the pharmacology of neonatal rat
RT glycine receptor subunit.*";
RL Neuron 5:867-873(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Spinal cord;
RX MEDLINE=91200276; PubMed=1707830;
RA Akagi H., Hirai K., Hishinuma F.;
RT "Cloning of a glycine receptor subtype expressed in rat brain and
RT spinal cord during a specific period of neuronal development.*";
RL FEBS Lett. 281:160-166(1991).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=91243883; PubMed=1645300;

```

```

RA Kuhse J., Kuratov A., Maulet Y., Malosio M.L., Schmieden V.,
RA Betz H.;
RT "Alternative splicing generates two isoforms of the alpha 2 subunit
RT of the inhibitory glycine receptor.*";
RL FEBS Lett. 283:73-77(1991).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Brain;
RA Beato M., Groot-Kormelink P.J., Colquhoun D., Sivillotti L.G.;
RT "Concentration dependence of single channel currents through rat
RT recombinant alpha 1 glycine receptors.*";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THE GLYCINE RECEPTOR IS A NEUROTRANSMITTER-GATED ION
CC CHANNEL. BINDING OF GLYCINE TO ITS RECEPTOR INCREASES THE CHLORIDE
CC CONDUCTANCE AND THUS PRODUCES HYPERPOLARIZATION (INHIBITION OF
CC NEURONAL FIRING).
CC -1- SUBUNIT: PENTAMER COMPOSED OF ALPHA (48 kDa) AND BETA (58 kDa)
CC SUBUNITS. A THIRD TYPE OF SUBUNIT (93 kDa) IS A PERIPHERAL
CC MEMBRANE PROTEIN ASSOCIATED WITH THE CYTOPLASMIC DOMAINS OF THE
CC RECEPTOR CORE.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: ALPHA-2* (SHOWN HERE) AND ALPHA-
CC 2B. ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: THE GLYCINE RECEPTOR IS ABUNDANT IN THE SPINAL
CC CORD AND BRAINSTEM OF VERTEBRATES.
CC -1- DEVELOPMENTAL STAGE: THE ALPHA-2* SUBUNIT ISOFORM IS PRESENT ONLY
CC IN NEONATAL RATS. ISOFORMS ALPHA-2A AND ALPHA-2B ARE PRESENT IN
CC THE PRE- AND NEONATAL BRAIN. AT LATER POSTNATAL STAGES, ALPHA-2A
CC LEVELS GREATLY DECREASE WHILE ALPHA-2B IS BARELY DETECTABLE.
CC -1- MISCELLANEOUS: THE ALPHA SUBUNIT BINDS STRYCHNINE.
CC -1- SUBSTITUTIONS: IDENTICAL TO THE HUMAN ALPHA-2 SUBUNIT, EXCEPT FOR
CC 5 SUBSTITUTIONS AT POSITIONS 18, 24, 37, 194 AND 404. SUBSTITUTION
CC AT POSITION 194 (G -> E) ACCOUNTS FOR THE LOWER STRYCHNINE
CC SENSITIVITY OBSERVED IN NEONATAL RATS.
CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: X57281; CAA40549.1; -
DR EMBL: X61159; CAA43471.1; -
DR EMBL: AJ310837; CAC35981.1; -
DR PIR: JN0112; JN0112.
DR PIR: S14816; S14816.
DR InterPro: IPR000188; GABAA_receptor.
DR InterPro: IPR001175; Neur_channel.
DR Pfam: PF02931; Neur_chan_LBD; 1.
DR Pfam: PF02932; Neur_chan_Memb; 1.
DR PRINTS: PRO0252; NRIONCHANNEL.
DR TIGRFAMs: TIGR00860; LIC; 1.
DR PROSITE: PS00236; NEURORTR_ION_CHANNEL; 1.
KW Receptor; Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
KW Transmembrane; Multigene family; Alternative splicing.
FT SIGNAL 1 27
FT CHAIN 28 452
FT DOMAIN 28 253
FT TRANSMEM 254 280
FT TRANSMEM 287 304
FT TRANSMEM 319 342
FT DOMAIN 343 423
FT TRANSMEM 424 441
FT SITE 194 194
FT DISULFID 172 186
FT DISULFID 232 243
FT CARBOHYD 72 72
FT CARBOHYD 103 103
FT VARSPLIC 85 86

```



FT VARIANT 194 194 E -> G.  
FT MUTAGEN 194 194 E->G: INCREASED SENSITIVITY TO  
STRYCHNINE.  
SQ SEQUENCE 452 AA: 52053 MW: A9BEC6D1552A3A1C CRC64:  
Query Match 74.2%; Score 1670; DB 1; Length 452;  
Best Local Similarity 75.3%; Pred. No. 9,3e-135;  
Matches 326; Conservative 32; Mismatches 49; Indels 26; Gaps 6;  
OY 1 MTTLPATSLFLLTLPGQVLLRVALAKE-EVKSQTGSGSPSPDFDLKMGRTSGVD 59  
DB 5 LVNITLAFAPFL-----GTNHFRAFCCKDHSRSGKHSQTLSDFLDKMGRTSGVD 59  
OY 60 ARIKRNKGPVNVTCNIFINSFSSVYKTTMDYRVNVELRQOMDPRLSYREYPPDSIDL 119  
DB 60 ARIKRNKGPVNVTCNIFINSFSSVYKTTMDYRVNVELRQOMDPRLSYREYPPDSIDL 119  
OY 120 DPMMDLSTWKPDLFFANKEGANFHEVTDNKLIRFKNGNVLYSTRLLTLLSCMDLKNF 179  
DB 120 DPMMDLSTWKPDLFFANKEGANFHEVTDNKLIRFKNGNVLYSTRLLTLLSCMDLKNF 179  
OY 180 PMDIQTCTMQLSESSILCSPLSLSVGYTKKDLVFEMLEDPAPVQVAGEGLTLPOFILR 239  
DB 180 PMDIQTCTMQLSESSILCSPLSLSVGYTKKDLVFEMLEDPAPVQVAGEGLTLPOFILR 239  
OY 240 DEKDLGCTKHYNTGKFTCIKVEKFLEROMGYLLIQMTIPSLIIVLSVSWFWINMDAP 299  
DB 240 DEKDLGCTKHYNTGKFTCIKVEKFLEROMGYLLIQMTIPSLIIVLSVSWFWINMDAP 299  
OY 225 EEKEIGYCTKHYNTGKFTCIKVEKFLEROMGYLLIQMTIPSLIIVLSVSWFWINMDAP 284  
DB 225 EEKEIGYCTKHYNTGKFTCIKVEKFLEROMGYLLIQMTIPSLIIVLSVSWFWINMDAP 284  
OY 300 ANVGIGITVLTMTQSSGRASLPKVSYSVKAIIDIMAAVCLFVRAALLEVAANFVSQ 359  
DB 285 ANVALGITTTLTMTQSSGRASLPKVSYSVKAIIDIMAAVCLFVRAALLEVAANFVSQ 344  
OY 360 HKEFTLRRLRRORLEEDIIQESREYFGYGLGHLQANRQDGPESGSIYS-POPPAPL 418  
DB 345 HKEFTLRRLRRORLEEDIIQESREYFGYGLGHLQANRQDGPESGSIYS-POPPAPL 401  
OY 419 LREGTTRKLYVD 431  
DB 402 -KDADAIRKKFVD 413  
RESULT 4  
GRA3\_HUMAN  
ID GRA3\_HUMAN STANDARD: PRT: 465 AA.  
AC 075311; 075816;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Glycine receptor alpha-3 chain precursor.  
GN GLRA3.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Fetal brain;  
RX MEDLINE=98344067; PubMed=9677400;  
RA Nikolic Z., Laube B., Weber R.G., Lichter P., Kioschis P., Poustka A.,  
Muehlhardt C., Becker C.-M.;  
RT "The human glycine receptor subunit alpha3. GLRA3 gene structure,  
chromosomal localization, and functional characterization of  
RT alternative transcripts.";  
RL J. Biol. Chem. 273:19708-19714(1998).  
CC -1- FUNCTION: THE GLYCINE RECEPTOR IS A NEUROTRANSMITTER-GATED ION  
CHANNEL. BINDING OF GLYCINE TO ITS RECEPTOR INCREASES THE CHLORIDE  
CONDUCTANCE AND THUS PRODUCES HYPERPOLARIZATION (INHIBITION OF  
NEURONAL FIRING).  
CC -1- SUBUNIT: PENTAMER COMPOSED OF LIGAND BINDING ALPHA (48 kDa) AND  
STRUCTURAL BETA (58 kDa) SUBUNITS. A THIRD TYPE OF SUBUNIT (93  
CC kDa) IS A PERIPHERAL MEMBRANE PROTEIN ASSOCIATED WITH THE  
CYTOPLASMIC DOMAINS OF THE RECEPTOR CORE.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; ALPHA3L (SHOWN HERE) AND  
CC ALPHA3K; ARE PRODUCED BY ALTERNATIVE SPLICING.  
CC -1- TISSUE SPECIFICITY: WIDELY DISTRIBUTED THROUGHOUT THE CENTRAL  
CC NERVOUS SYSTEM.  
CC -1- DOMAIN: THE N-TERMINAL DOMAIN CARRIES STRUCTURAL DETERMINANTS  
CC ESSENTIAL FOR AGONIST AND ANTAGONIST BINDING. TM2 IS THOUGHT TO  
CC FORM THE INNER WALL OF THE CHLORIDE CHANNEL. THE CYTOPLASMIC LOOP  
CC IS AN IMPORTANT DETERMINANT OF CHANNEL INACTIVATION KINETICS.  
CC -1- MISCELLANEOUS: THE ALPHA SUBUNIT BINDS STRYCHNINE.  
CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC  
CC EMBL: AF017724; AAC39919.1; JOINED.  
CC EMBL: AF017715; AAC39919.1; JOINED.  
CC EMBL: AF017716; AAC39919.1; JOINED.  
CC EMBL: AF017717; AAC39919.1; JOINED.  
CC EMBL: AF017718; AAC39919.1; JOINED.  
CC EMBL: AF017719; AAC39919.1; JOINED.  
CC EMBL: AF017720; AAC39919.1; JOINED.  
CC EMBL: AF017721; AAC39919.1; JOINED.  
CC EMBL: AF017722; AAC39919.1; JOINED.  
CC EMBL: AF017723; AAC39919.1; JOINED.  
CC EMBL: U93917; AAC39917.1; JOINED.  
CC GeneW: HGNC:4328; GLRA3.  
CC MIM: 600421;  
CC InterPro: IPR000188; GABA\_A\_receptor.  
CC InterPro: IPR001175; Neur\_chan.  
CC Pfam: PF02932; Neur\_chan\_LBD; 1.  
CC DR PFAM: PF02932; Neur\_chan\_LBD; 1.  
CC DR TIGR: TIGR00860; LIC; 1.  
CC DR PROSITE: PS00236; NEURORTR\_ION\_CHANNEL; 1.  
CC KW Receptor; Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;  
CC Transmembrane; Multigene family; Alternative splicing.  
CC  
CC FT SIGNAL 1 33  
CC FT CHAIN 34 465 GLYCINE RECEPTOR ALPHA-3 CHAIN.  
CC FT DOMAIN 251 251 EXTRACELLULAR (PROBABLE).  
CC FT TRANSMEM 253 279 PROBABLE.  
CC FT TRANSMEM 286 303 PROBABLE.  
CC FT TRANSMEM 315 341 PROBABLE.  
CC FT DOMAIN 342 433 CYTOPLASMIC (PROBABLE).  
CC FT TRANSMEM 434 451 PROBABLE.  
CC FT DISULFID 171 185 BY SIMILARITY.  
CC FT DISULFID 231 242 BY SIMILARITY.  
CC FT CARBOHYD 71 71 N-LINKED (GLCNAC... ) (POTENTIAL).  
CC FT VARSPPLIC 358 372 MISSING (IN ISOFORM ALPHA3K).  
CC FT CONFLICT 461 461 MISSING (IN REF. 1; AAC39917).  
CC SQ SEQUENCE 465 AA: 53937 MW: B8A267A310C6E64 CRC64:  
Query Match 71.2%; Score 1602.5; DB 1; Length 465;  
Best Local Similarity 73.8%; Pred. No. 5,5e-129;  
Matches 312; Conservative 35; Mismatches 43; Indels 33; Gaps 6;  
OY 22 LLRVALAKEEVKSGTKSGSPSPDFDLKMGRTSGYDARIRPNKGPVNVTCNIFINS 81  
DB 21 LLLSLVATKETDSARSAPSDFDLKMGRTSGYDARIRPNKGPVNVTCNIFINS 80  
OY 82 ESSVYKTTMDYRVNVLFRQOMDPRLSYREYPPDSIDLDPMSLDSIMKPDFFANKEGAN 141  
DB 81 FIEVTTDNKLIRFKNGNVLYSTRLLTLLSCMDLKNFPMDDIQTCTMQLSESSILCSPLP 201  
OY 142 FIEVTTDNKLIRFKNGNVLYSTRLLTLLSCMDLKNFPMDDIQTCTMQLSESSILCSPLP 201  
DB 141 FIEVTTDNKLIRFKNGNVLYSTRLLTLLSCMDLKNFPMDDIQTCTMQLSESSILCSPLP 190









DB 409 PEMRKLFT 417

RESULT 8

GRAL\_BRARE STANDARD: PRT: 444 AA.

AC 093430:

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Glycine receptor, alpha1 chain precursor.

GN GLRA1.

OS Brachydanio rerio (Zebrafish) (Danio rerio).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.

OC NCBL\_TaxId=7955;

OX NCBL\_TaxId=7955;

RN (1)

RN SEQUENCE FROM N.A.

RC TISSUE-Brain;

RX MEDLINE-99202839; PubMed-10189956;

RA David-Mattine B., Goblet C., de Saint Jan D., Fucile S., Devignot V., Bregestovski P., Korn H.;

RT Cloning, expression and electrophysiological characterization of glycine receptor alpha subunit from zebrafish.";

RL Neuroscience 90:303-317(1999).

CC -1- FUNCTION: THE GLYCINE RECEPTOR IS A NEUROTRANSMITTER-GATED ION CHANNEL. BINDING OF GLYCINE TO ITS RECEPTOR INCREASES THE CHLORIDE CONDUCTANCE AND THUS PRODUCES HYPERPOLARIZATION (INHIBITION OF NEURONAL FIRING).

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.

CC -1- TISSUE SPECIFICITY: EXPRESSED IN BRAIN.

CC -1- MISCELLANEOUS: THIS RECEPTOR IS ACTIVATED BY GLYCINE AND ANTAGONIZED BY STRYCHNINE. CAN ALSO BE ACTIVATED BY GABA AND INHIBITED BY BICUCULLINE.

CC -1- MISCELLANEOUS: HIGHLY SENSITIVE TO ACTIVATION BY TAURINE DESPITE THE PRESENCE OF A VALINE IN POSITION 135. IN MAMMALS VALINE AT THIS POSITION CAUSES A DRASTIC LOSS OF TAURINE EFFICACY.

CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC -----

DR EMBL: AJ005812; CA06711.1; -;

DR ZFIN: ZDB-GENE-991117-1; glra1.

DR InterPro: IPR000188; GABA\_A\_receptor.

DR InterPro: IPR001175; Neur\_channel.

DR Pfam: PF02931; Neur\_chan\_LBD; 1.

DR Pfam: PF02932; Neur\_chan\_memb; 1.

DR PRINTS: PR00252; NRIONCHANNEL.

DR TIGR: TIGR00860; LIC; 1.

DR PROSITE: PS00236; NEURON\_CHANNEL; 1.

DR Receptor: Postsynaptic membrane; Ionic channel; Glycoprotein; Signal; Transmembrane; Multigene family.

KW SIGNAL

FT 1 24

FT CHAIN

FT 25 444

FT DOMAIN

FT 25 444

FT TRANSMEM

FT 248 268

FT TRANSMEM

FT 277 294

FT TRANSMEM

FT 311 331

FT DOMAIN

FT 332 415

FT TRANSMEM

FT 416 436

FT DISULFID

FT 162 176

FT DISULFID

FT 222 233

FT CARBOHYD

FT 62 62

FT BINDING

FT 221 221

FT BINDING

FT 226 226

SQ SEQUENCE 444 AA: 50765 MW: 7756FC8C14E80581 CRC64:

Query Match 69.8%; Score 1573; DB 1: Length 444;

Best Local Similarity 71.9%; Pred. No. 1.7e-126;

Matches 302; Conservative 43; Mismatches 33; Indels 22; Gaps 5;

QY 12 LLLWTLPGVLLVLAKEEVKSGTSGSQSPSDFLLKMGRTGYDARIRPNFKGPPV 71

DB 6 IYLM----ETIVFSLASQQAARKAASPMPSEFLDKMGKSGYDARIRPNFKGPPV 61

QY 72 NVTGNTFINSFSSVYKTTMDYRVNVLFRQMDPRLSTRVYDDSLDPSMLDSIKRPD 131

DB 62 NVTCNIFINSFGSIAETMDYRVNIFLRQMDPRLAYSEVPDDSLDPSMLDSIKRPD 121

QY 132 LFFANDEKGNFHEVTTDNKLRIFFKNGVLYSIRTLTLCSCMDLKNPMDIQGTWOLE 191

DB 122 LFFANDEKGNFHEVTTDNKLRISSKNGVLYSIRTLTLCSCMDLKNPMDIQGTWOLE 181

QY 192 SSSILCSPLPSLSLSGYTKMDLVEFMELEDAFVAVAGTLTLPQFILLDEKDLGCTKHY 251

DB 182 -----SFGYTMNDLIFEW--DEKGAOVADGTLTLPQFILLDEKDLGCTKHY 226

QY 252 NTGKFTCLFVKHLEBROKGYLIQMYTSLIYLISWVFWIMDAPARGLITVLT 311

DB 227 NTGKFTCLFARHLEBROKGYLIQMYTSLIYLISWVFWIMDAPARGLITVLT 286

QY 312 MTTGSSGRASLPKYSYKATIDIMAVCLLPFAALLFVAALNPFNSRQHKFIRLROR 371

DB 287 MTTGSSGRASLPKYSYKATIDIMAVCLLPFAALLFVAALNPFNSRQHKFIRLROR --R 344

QY 372 RQRLLEDILQESRFYFEGYGLG-HCLQARDGGPMESGTSYSPQPPALRREGETTRKLYV 430

DB 345 RRLKEDDEGDGRFSFAAVGMGPACLOAKDGMAIKGNMNNAPTSTNPPEKTYEEMRKLEI 404

RESULT 9

GRAL\_MOUSE STANDARD: PRT: 457 AA.

AC 064018; Q64019; Q9R0Y6; Q9R0Y7;

DT 01-NOV-1997 (Rel. 35, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Glycine receptor alpha-1 chain precursor (Glycine receptor 48 kDa subunit) (Strychnine binding subunit).

GN GLRA1.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OC NCBL\_TaxId=10090;

RN (1)

RN SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND VARIANT SPD SER-80.

RX MEDLINE-95004575; PubMed-7920629;

RA Ryan S.G., Buckwalter M.S., Lynch J.W., Handford C.A., Segura L., Shiang R., Wasmuth J.J., Camper S.A., Schofield P., O'Connell P.;

RT A missense mutation in the gene encoding the alpha 1 subunit of the inhibitory glycine receptor in the spasmodic mouse.";

RL Nat. Genet. 7:131-135(1994).

RN [2]

RN SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.

RC STRAIN-BALB/C;

RX MEDLINE-94132024; PubMed-7507926;

RA Matzenbach B., Maulet Y., Sefton L., Courtlier B., Avner P., Guenet J.L., Betz H.;

RT "Structural analysis of mouse glycine receptor alpha subunit genes. Identification and chromosomal localization of a novel variant.";

RL J. Biol. Chem. 269:2607-2612(1994).

CC -1- FUNCTION: THE GLYCINE RECEPTOR IS A NEUROTRANSMITTER-GATED ION CHANNEL. BINDING OF GLYCINE TO ITS RECEPTOR INCREASES THE CHLORIDE CONDUCTANCE AND THUS PRODUCES HYPERPOLARIZATION (INHIBITION OF NEURONAL FIRING).

CC -1- SUBUNIT: PENTAMER COMPOSED OF ALPHA (48 KDA) AND BETA (58 KDA) SUBUNITS. A THIRD TYPE OF SUBUNIT (93 KDA) IS A PERIPHERAL MEMBRANE PROTEIN ASSOCIATED WITH THE CYTOPLASMIC DOMAINS OF THE

CC RECEPTOR CORE.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A LONG FORM (SHOWN HERE) AND A  
 CC SHORT FORM: ARE PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- DISEASE: DEFECTS IN GIRAL ARE A CAUSE OF SPASMODIC (SPD), A MOUSE  
 CC MUTANT WHICH RESEMBLES TO THE HUMAN NEUROLOGICAL DISEASE,  
 CC HYPERKALEMIA (OR STARTLE DISEASE (STHE)).  
 CC -1- MISCELLANEOUS: THE ALPHA SUBUNIT BINDS STRYCHNINE.  
 CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: S73717; AAB32157.2; -  
 CC EMBL: S73718; AAB32158.2; -  
 CC EMBL: X75832; CAB52398.1; -  
 CC EMBL: X75833; CAB52398.1; JOINED.  
 CC EMBL: X75834; CAB52398.1; JOINED.  
 CC EMBL: X75835; CAB52398.1; JOINED.  
 CC EMBL: X75836; CAB52398.1; JOINED.  
 CC EMBL: X75837; CAB52398.1; JOINED.  
 CC EMBL: X75838; CAB52398.1; JOINED.  
 CC EMBL: X75839; CAB52398.1; JOINED.  
 CC EMBL: X75840; CAB52398.1; JOINED.  
 CC EMBL: X75832; CAB52399.1; -  
 CC EMBL: X75833; CAB52399.1; JOINED.  
 CC EMBL: X75834; CAB52399.1; JOINED.  
 CC EMBL: X75835; CAB52399.1; JOINED.  
 CC EMBL: X75836; CAB52399.1; JOINED.  
 CC EMBL: X75837; CAB52399.1; JOINED.  
 CC EMBL: X75838; CAB52399.1; JOINED.  
 CC EMBL: X75839; CAB52399.1; JOINED.  
 CC EMBL: X75840; CAB52399.1; JOINED.  
 CC MGD: MGI:95747; Glral.  
 CC Interpro: IPR000188; GABA\_receptor.  
 CC Interpro: IPR001175; Neur\_channel.  
 CC Pfam: PF02931; Neur\_chan\_LBD.1.  
 CC Pfam: PF02932; Neur\_chan\_memb.1.  
 CC PRINTS: PR00252; NRIONCHANNEL.  
 CC TIGRFAMs: TIGR00860; LIC.1.  
 CC PROSITE: PS00236; NEUOTR\_ION\_CHANNEL.1.  
 CC Receptor: Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;  
 CC Transmembrane; Multigene family; Alternative splicing;  
 CC Disease mutation.  
 CC KW SIGNAL.1  
 CC FT CHAIN 1 28  
 CC FT DOMAIN 29 457 GLYCINE RECEPTOR ALPHA-1 CHAIN.  
 CC FT TRANSMEM 248 274 EXTRACELLULAR (PROBABLE).  
 CC FT TRANSMEM 281 298 PROBABLE.  
 CC FT TRANSMEM 313 336 PROBABLE.  
 CC FT DOMAIN 337 420 CYTOPLASMIC (PROBABLE).  
 CC FT TRANSMEM 429 446 PROBABLE.  
 CC FT DISULFID 166 180 BY SIMILARITY.  
 CC FT DISULFID 226 237 BY SIMILARITY.  
 CC FT CARBOHYD 66 66 N-LINKED (GLUCAC. .) (PROBABLE).  
 CC FT BINDING 225 225 STRYCHNINE (OR 220) (BY SIMILARITY).  
 CC FT BINDING 230 230 STRYCHNINE (OR 225) (BY SIMILARITY).  
 CC FT VARSPLIC 354 361 MISSING (IN SHORT ISOFORM).  
 CC FT VARIANT 80 84 A -> S (IN SPD).  
 CC FT CONFLICT 84 84 M -> I (IN REF. 2).  
 CC FT CONFLICT 426 429 ISRI -> NISH (IN REF. 2).  
 CC FT SEQUENCE 457 AA; 52656 MW; 29268DC4991A6E20 CRC64;  
 CC Query Match 69.7%; Score 1569.5; DB 1; Length 457;  
 CC Best Local Similarity 71.5%; Pred. No. 3.5e-126;  
 CC Matches 309; Conservative 40; Mismatches 54; Indels 29; Gaps 8;  
 CC Oy 8 TLSFLMLMTLPBGVLLRVALLAKEEVKSGTQSGQSPMSDFDLKLMGRTSGYDARIRPNFK 67

Db 6 TLRFYLMETI---VFESLA-ASKEAEARSAAPKPMSPSDFDLKLMGRTSGYDARIRPNFK 61  
 Oy 68 GPPVNTVNTCFNFSFSSVYTTTMDYRNVNVLROOANDPRLSYKEYPPDSDLDPMSLDST 127  
 Db 62 GPPVNVSCNFNFSFSGIAETTDYRNVNIFLROOANDPRLAYNEYPPDSDLDPMSLDST 121  
 Oy 128 WKPDLPFANEGKGFNFHVTTDNKLRIFKNGVNLVSTRLLISCLMDLKNFPMQIQCTP 187  
 Db 122 WKPDLPFANEGKGFNFHVTTDNKLRIFKNGVNLVSTRLLISCLMDLKNFPMQIQCTP 181  
 Oy 188 MOLESSIILCSPLSLSVGYTWMKDLVFEMLEDAVAVQVAGLTPQFTLRDEKLGCC 247  
 Db 182 MOLE-----SEGYTMNDLIFEMOEG-AYQVADSLTLPOFTLKEEKDLRYC 226  
 Oy 248 TKHYNTGKFTCEYKFFLEPROMGYLYLQNYIPSLLYIISWFSWIMDAAPRVGIGT 307  
 Db 227 TKHYNTGKFTCEYKFFLEPROMGYLYLQNYIPSLLYIISWFSWIMDAAPRVGIGT 286  
 Oy 308 TVLMTTQSSGRASLPKVSIVKAIIDIMAVCLLFEPAALLEYAAINFSRQKKEFTLR 367  
 Db 287 TVLMTTQSSGRASLPKVSIVKAIIDIMAVCLLFEPAALLEYAAINFSRQKKEFTLR 346  
 Oy 368 RROPROR-----LEEDIIDSRFFRYGGLG-HCLQARDGGMESGIGTSP--OPAPL 418  
 Db 347 RKRHRHKSPLMLNLFQDDEGEGGRFNFSAYGMGPACLOAKDGI SVKGANNNNTTNPAP- 405  
 Oy 419 LREGETRKLIV 430  
 Db 406 SKSPEEMRKLFI 417

RESULT 10  
 GRAL\_RAT STANDARD; PRT; 457 AA.  
 ID GRAL\_RAT  
 AC P07727;  
 DT 01-APR-1988 (Rel. 07, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Glycine receptor alpha-1 chain precursor (Glycine receptor 48 kDa subunit) (Strychnine binding subunit).  
 GN GRAL OR GLYR.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxId=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A. (SHORT ISOFORM).  
 RX MEDLINE=91367372; Pubmed=1716350;  
 RA Akagi H., Hirai K., Hishinuma F.;  
 RT "Functional properties of strychnine-sensitive glycine receptors expressed in *Xenopus* oocytes injected with a single mRNA.";  
 RL Neurosci. Res. 11:28-40(1991).  
 RN [2]  
 RP SEQUENCE OF 23-457 FROM N.A. (SHORT ISOFORM), AND SEQUENCE OF 29-49. STRAIN=Wistar;  
 RC MEDLINE=87258250; Pubmed=3037383;  
 RA Grenningloh G., Rindler A., Schmitt B., Methfessel C., Zensen M., Beyreuther K., Gundelfinger E.D., Betz H.;  
 RT "The strychnine-binding subunit of the glycine receptor shows homology with nicotinic acetylcholine receptors.";  
 RL Nature 328:215-220(1987).  
 RN [3]  
 RP SEQUENCE OF 1-30 FROM N.A.  
 RX MEDLINE=90180468; Pubmed=2483325;  
 RA Sontelmeier H., Becker C.M., Pritchett D.B., Schofield P.R., Grenningloh G., Kettenmann H., Betz H., Seeburg P.H.;  
 RT "Functional chloride channels by mammalian cell expression of rat glycine receptor subunit.";  
 RL Neuron 2:1491-1497(1989).  
 RN [4]  
 RP STRYCHNINE BINDING SITE.  
 RX MEDLINE=91027731; Pubmed=2171639;

RA Ruiz-Gomez A., Morato E., Garcia-Calvo M., Valdivieso F.,  
 RA Mayor F. Jr.;  
 RT "Localization of the strychnine binding site on the 48-kilodalton  
 RT subunit of the glycine receptor.";  
 RL Biochemistry 29:7033-7040(1990).  
 RN (5)  
 RP ALTERNATIVE SPLICING.  
 RC TISSUE-Spinal cord;  
 RX MEDLINE-9115810; PubMed-1703526;  
 RA Malosio M.L., Greeningloh G., Kuhse J., Schmieden V., Schmitt B.,  
 RA Prior P., Betz H.;  
 RT "Alternative splicing generates two variants of the alpha 1 subunit  
 RT of the inhibitory glycine receptor.";  
 RL J. Biol. Chem. 266:2048-2053(1991).  
 CC -1- FUNCTION: THE GLYCINE RECEPTOR IS A NEUROTRANSMITTER-GATED ION  
 CC CHANNEL. BINDING OF GLYCINE TO ITS RECEPTOR INCREASES THE CHLORIDE  
 CC CONDUCTANCE AND THUS PRODUCES HYPERPOLARIZATION (INHIBITION OF  
 CC NEURONAL FIRING).  
 CC -1- SUBUNIT: PENTAMER COMPOSED OF ALPHA (48 kDa) AND BETA (58 kDa)  
 CC SUBUNITS. A THIRD TYPE OF SUBUNIT (93 kDa) IS A PERIPHERAL  
 CC MEMBRANE PROTEIN ASSOCIATED WITH THE CYTOPLASMIC DOMAINS OF THE  
 CC RECEPTOR CORE.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A LONG FORM (SHOWN HERE) AND A  
 CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- MISCELLANEOUS: THE ALPHA SUBUNIT BINDS STRYCHNINE.  
 CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC  
 CC EMBL: D00833; BA00707.1; -  
 CC EMBL: Y00276; CAA68378.1; -  
 CC EMBL: M63915; AAB63490.1; -  
 CC EMBL: X55246; CAA38987.1; -  
 CC PIR: A27141; A27141.  
 CC PIR: A38597; A38597.  
 CC PIR: JN0014; JN0014.  
 CC InterPro: IPR000188; GABA\_A\_receptor.  
 CC InterPro: IPR001175; Neur\_chan.  
 CC Pfam: PF02931; Neur\_chan\_LBD; 1.  
 CC Pfam: PF02932; Neur\_chan\_memb; 1.  
 CC PRINTS: PR00252; NRIONCHANNEL.  
 CC TIGRFS: TIGR00860; LIC; 1.  
 CC PROSITE: PS00236; NEUROTR\_ION\_CHANNEL; 1.  
 CC Receptor: Postsynaptic membrane; Ionic channel; Glycoprotein; signal;  
 CC Transmembrane; Multigene family; Alternative splicing.  
 CC  
 CC FT CHAIN 1 457 GLYCINE RECEPTOR ALPHA-1 CHAIN.  
 CC FT DOMAIN 29 247 EXTRACELLULAR (PROBABLE).  
 CC FT TRANSMEM 248 274 PROBABLE.  
 CC FT TRANSMEM 281 298 PROBABLE.  
 CC FT TRANSMEM 313 336 PROBABLE.  
 CC FT DOMAIN 337 420 CYTOPLASMIC (PROBABLE).  
 CC FT TRANSMEM 421 445 PROBABLE.  
 CC FT DISULFID 166 180 BY SIMILARITY.  
 CC FT DISULFID 226 237 BY SIMILARITY.  
 CC FT CARBOHYD 66 66 N-LINKED (GLCNAC. . .) (PROBABLE).  
 CC FT BINDING 225 225 STRYCHNINE (OR 230).  
 CC FT BINDING 230 230 STRYCHNINE (OR 225).  
 CC FT VARSPPLIC 354 361 MISSING (IN SHORT ISOFORM).  
 CC FT CONFLICT 1 22 MISENTRELYMETIVFEFLAA -> MGKSPGLDLYLWMT  
 CC FT CONFLICT 22 LEFP (IN REF. 3).  
 CC FT CONFLICT 26 A -> V (IN REF. 2).  
 CC FT SEQUENCE 457 AA; 52616 MW; 0379C33A83C71FAA CRC64;  
 CC  
 CC Query Match 69.68; Score 1568.5; DB 1; Length 457;  
 CC Best Local Similarity 71.58; Pred. No. 4.2e-126;

Matches 308; Conservative 39; Mismatches 57; Indels 27; Gaps 7;  
 QY 8 TLSELLLTLPQGVLLRYALAKEEYKSGTCKSOPSPDLDKLMGRSYDARIRPEK 67  
 DB 6 TLREYLMETI---VFSSIA-ASKEDAAKRSAPKPPSPDFDKLMDKGRISGTDARIRPEK 61  
 QY 68 GPPVAVTCNIFINSSFYTKTMDYRVNVFLRQOQNDPRLSYREXPDDSLDLPDSMLDSI 127  
 DB 62 GPPVAVSCNIFINSSGSLAETTMQDVRVNIPLRQOQNDPRLVYNEPDDSLDLPDSMLDSI 121  
 QY 128 WKPDLPFANEGANHEVTTDKLIRKNGNVLYSLTLTLCMDLAKFPMDDIQTCT 187  
 DB 122 WKPDLPFANEGANHEVTTDKLIRKNGNVLYSLTLTLCMPDKNFPMDDVQTCI 181  
 QY 188 MOLESSSLCSPLSLSVGYTKMDIVFEMLEDAPAQVQEGSLTLPQFIRDRKDGCC 247  
 DB 182 MOLE-----SFGITMDLDFEPOEG-AQVADGTLTLPQFIRKEKDLRYC 226  
 QY 248 TKHYVTGKFTCLIEVNFHLEROMGYLLQMYIPSLDIVILTSWSEFWINNDAPARVGLGIT 307  
 DB 227 TKHYVTGKFTCLIEARFHLERQGYLLQMYIPSLDIVILTSWSEFWINNDAPARVGLGIT 286  
 QY 308 TVLTWTQSSGSRASLPVSYVYKATDIMAVCLIFVFAALLEYAINFVSROHKEFTLR 367  
 DB 287 TVLTWTQSSGSRASLPVSYVYKATDIMAVCLIFVFAALLEYAINFVSROHKEFTLR 346  
 QY 368 RRORROR-----LEEDITQESREFYRGYGLG-HCLQARDGPGMGSGSYVPOPPAPL-L 419  
 DB 347 RRRRHKKRPMNLFPQDDGEGEGRFNFSAVGMGPRACLOAKDGLSVGANNNTTNPAPAPS 406  
 QY 420 REGETTRKLYV 430  
 DB 407 KSPDEMRKLF 417  
 RESULT 11  
 GRB\_MOUSE  
 ID GRB\_MOUSE STANDARD; PRT; 496 AA.  
 AC P48168;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Glycine receptor beta chain precursor (glycine receptor 58 kDa  
 DE subunit).  
 GN GLRB.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=95004576; PubMed=7920630;  
 RA Kingsmore S.F., Giros B., Suh D., Bieniarz M., Caron M.G.,  
 RA Seidlin M.F.;  
 RT "Glycine receptor beta-subunit gene mutation in spastic mouse  
 RT associated with LINE-1 element insertion.";  
 RL Nat. Genet. 7:136-141(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A. AND VARIANTS SPASTIC.  
 RC STRAIN=C57BL/6, and BALB/c; TISSUE=Brain, and Liver;  
 RX MEDLINE=9503198; PubMed=7946325;  
 RA Muehlhardt C., Fischer M., Gass P., Simon-Chazottes D.,  
 RA Guenet J.-L., Kuhse J., Betz H., Becker C.M.;  
 RT "The spastic mouse: aberrant splicing of glycine receptor beta  
 RT subunit mRNA caused by intronic insertion of L1 element.";  
 RL Neuron 13:1003-1015(1994).  
 CC -1- FUNCTION: THE GLYCINE RECEPTOR IS A NEUROTRANSMITTER-GATED ION  
 CC CHANNEL. BINDING OF GLYCINE TO ITS RECEPTOR INCREASES THE CHLORIDE  
 CC CONDUCTANCE AND THUS PRODUCES HYPERPOLARIZATION (INHIBITION OF  
 CC NEURONAL FIRING).  
 CC -1- SUBUNIT: PENTAMER COMPOSED OF ALPHA AND BETA SUBUNITS.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.





Db 63 LL---VSYDPRIRPNFKGIPVDVYVNFINSFGSIQETTMQRYVNIPLRQKNDPRLKLP 119  
 QY 108 -SYREYPPDSLDLDPMSIMDSIMKPDLEFANEKANFEHVTITDNKLLRIFKNGVLYSIRL 166  
 Db 120 SDFR--GSDALVTDPMTWKCLMKPDLFEFANEKANFEHVTIOENILLFFIRGQDVLVSMRL 177  
 QY 167 TLILSCMLDKNFPMDIQCTMQOLESSILCSPLPSLSVGTVMKDLVFEWLEDPAPVQ 226  
 Db 178 SITLSCPLDLTLFPMDTQRCKMQLE-----SFGYTTDDLRFIMQSGDP-VQ 222  
 QY 227 VAEGTLTLPF--ILRDEKDLGCTKHY--NTGKFTCIEVKFHLERQMGYLLQMTYIPSLIIV 284  
 Db 223 L-EKIALPQFDIKEDIEYNGCTKYKGTGYTCVEIFTLRQVGYFMGVYAPFLIIV 281  
 QY 285 ILSWSEFIMNDAPARVGLGITVLTMTTOSGSRASLPKVSYSYKADIMAVCLLFEV 344  
 Db 282 VLSWSEFIMNDASARVPLGIFSVLSLASECTTLAELPKVSYSYKADIMAVCLLFGF 341  
 QY 345 AALLEYAAI---NFVSRQKKEFLRLRROR 371  
 Db 342 ASLVEYAVVQVLMNPKRVEAEKARIKAEQ 372

## RESULT 13

GRB\_HUMAN STANDARD: PRT: 497 AA.  
 AC P48167;

DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Glycine receptor beta chain precursor (Glycine receptor 58 kDa subunit).  
 GN GLRB.

OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 NC NCB1\_TaxID=9606;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Hippocampus;  
 RX MEDLINE=96352561; PubMed=8717357;

RA Handford C.A., Lynch J.W., Baker E., Webb G.C., Ford J.H., Sutherland G.R., Schofield P.R.;  
 RT "The human glycine receptor beta subunit: primary structure, functional characterisation and chromosomal localisation of the human and murine genes";  
 RT Brain Res. Mol. Brain Res. 35:211-219(1996).  
 RL [2]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=96341117; PubMed=9676428;  
 RA Milani N., Meilhardt C., Weber R.G., Lichter P., Kioschis P., Pouska A., Becker C.-M.;  
 RT "The human glycine receptor beta subunit gene (GLRB): structure, refined chromosomal localization, and population polymorphism";  
 RL Genomics 50:341-345(1998).

CC -1- FUNCTION: THE GLYCINE RECEPTOR IS A NEUROTRANSMITTER-GATED ION CHANNEL. BINDING OF GLYCINE TO ITS RECEPTOR INCREASES THE CHLORIDE CONDUCTANCE AND THUS PRODUCES HYPERPOLARIZATION (INHIBITION OF NEURONAL FIRING).  
 CC -1- SUBUNIT: PENTAMER COMPOSED OF ALPHA AND BETA SUBUNITS.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC EMBL: U33267; AAB37750.1;  
 CC EMBL: AF094754; AAC71033.1;

DR EMBL: AF094755; AAC71034.1; -.  
 DR Genes; HGNC:4329; GLRB.  
 DR MIM: 138492; -.  
 DR InterPro: IPR000188; GABA\_receptor.  
 DR InterPro: IPR001175; Neur\_channel.  
 DR Pfam: PF02931; Neur\_chan\_LBD; 1.  
 DR Pfam: PF02932; Neur\_chan\_membr; 1.  
 DR PRINTS: PR00252; NRIONCHANNEL.  
 DR TIGRFAMs: TIGR00860; LIC; 1.  
 DR PROSITE: PS00236; NEUROR\_ION\_CHANNEL; 1.  
 DR Receptor; Postsynaptic membrane; Ionic channel; Glycoprotein; Signal; Transmembrane.  
 KM SIGNAL 1 22  
 FT CHAIN 23 497  
 FT DOMAIN 23 265  
 FT TRANSMEM 266 290  
 FT TRANSMEM 299 316  
 FT TRANSMEM 331 354  
 FT DOMAIN 355 478  
 FT TRANSMEM 479 496  
 FT DISULFID 183 197  
 FT CARBOHYD 54 54  
 FT CARBOHYD 242 242  
 SQ SEQUENCE 497 AA: 56122 MW: 610840760103625 CRC64;

Query Match 38.0%; Score 855.5; DB 1; Length 497;  
 Best Local Similarity 47.4%; Pred. No. 2.7e-65;  
 Matches 186; Conservative 64; Mismatches 89; Indels 53; Gaps 11;

QY 8 TLSELLLTLPDGOVILRALAKEVSKGSKQSQPA-----SPSDFLD 49  
 Db 6 TTAFLILSL---WVEEAYSKEKSSKKKKQYOLCPSSQSAEDLARVPANSTSLN 61  
 QY 50 KIMGRTSGYDARIRPNFKGPPVNYTCNIFINSFSSVTKTMDYRVNVEFLRQWMDPRL- 107  
 Db 62 RLL---VSYDPRIRPNFKGIPVDVYVNFINSFGSIQETTMQRYVNIPLRQKNDPRLK 118  
 QY 108 -SYREYPPDSLDLDPMSIMDSIMKPDLEFANEKANFEHVTITDNKLLRIFKNGVLYSIR 165  
 Db 119 PSDFR--GSDALVTDPMTWKCLMKPDLFEFANEKANFEHVTIOENILLFFIRGQDVLVSMR 176  
 QY 166 LTLILSCMLDKNFPMDIQCTMQOLESSILCSPLPSLSVGTVMKDLVFEWLEDPAPV 225  
 Db 177 SITLSCPLDLTLFPMDTQRCKMQLE-----SFGYTTDDLRFIMQSGDP-V 221  
 QY 226 VAEGTLTLPF--ILRDEKDLGCTKHY--NTGKFTCIEVKFHLERQMGYLLQMTYIPSLI 283  
 Db 222 OL-EKIALPQFDIKEDIEYNGCTKYKGTGYTCVEIFTLRQVGYFMGVYAPFLI 280  
 QY 284 VLSWSEFIMNDAPARVGLGITVLTMTTOSGSRASLPKVSYSYKADIMAVCLLFEV 343  
 Db 281 VLSWSEFIMNDASARVPLGIFSVLSLASECTTLAELPKVSYSYKADIMAVCLLFG 340  
 QY 344 AALLEYAAI---NFVSRQKKEFLRLRROR 371  
 Db 341 ASLVEYAVVQVLMNPKRVEAEKARIKAEQ 372

## RESULT 14

GAB\_LYMST STANDARD: PRT: 499 AA.  
 AC P26714;

DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Gamma-aminobutyric-acid receptor beta subunit precursor (GABA(A) receptor).  
 OS Lymnaea stagnalis (Great pond snail).  
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;  
 CC Lymnaeidae; Lymnaea.

NC NCB1\_TaxID=6523;  
 RX [1]  
 RP SEQUENCE FROM N.A.



RX MEDLINE-92007774; PubMed-1655414;  
 RA Harvey R.J., Vreugdenhil E., Zaman S.H., Bhandal N.S.,  
 RA Sharwood P.N.R., Barnard E.A., Darlison M.G.;  
 RT "Sequence of a functional invertebrate GABA<sub>A</sub> receptor subunit which  
 can form a chimeric receptor with a vertebrate alpha subunit.";  
 RL EMBO J. 10:3239-3245(1991).  
 CC -1- FUNCTION: GABA, THE MAJOR INHIBITORY NEUROTRANSMITTER IN THE  
 CC VERTEBRATE BRAIN, MEDIATES NEURONAL INHIBITION BY BINDING TO THE  
 CC GABA/BENZODIAZEPINE RECEPTOR AND OPENING AN INTEGRAL CHLORIDE  
 CC CHANNEL.  
 CC -1- SUBUNIT: GENERALLY PENTAMERIC. THERE ARE FIVE TYPES OF GABA(A)  
 CC RECEPTOR CHAINS: ALPHA, BETA, GAMMA, DELTA, AND RHO.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: X58638; CAA41495.1; -  
 CC PIR: S17785; S17785.  
 DR InterPro: IPR000188; GABA<sub>A</sub>receptor.  
 DR InterPro: IPR001175; Neur\_chan.  
 DR Pfam: PF02931; Neur\_chan\_LBD; 1.  
 DR Pfam: PF02932; Neur\_chan\_memb; 1.  
 DR PRINTS: PR00252; NRIONCHANNEL.  
 DR TIGRFS: TIGR00860; LIC; 1.  
 DR PROSITE: PS00236; NEUROR\_ION\_CHANNEL; 1.  
 DR Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;  
 KM Transmembrane; Multigene family;  
 FT SIGNAL 1 23  
 FT CHAIN 24 499  
 FT DOMAIN 24 251  
 FT TRANSMEM 252 273  
 FT TRANSMEM 278 299  
 FT TRANSMEM 311 333  
 FT DOMAIN 334 475  
 FT TRANSMEM 476 499  
 FT CARBOHYD 32 32  
 FT CARBOHYD 98 98  
 FT CARBOHYD 106 106  
 FT CARBOHYD 152 152  
 FT DISULFID 167 181  
 SQ SEQUENCE 499 AA; 57081 MW; F81C9F2A850D62F4 CRC64;  
 Query Match 32.9%; Score 742; DB 1; Length 499;  
 Best Local Similarity 42.2%; Pred. No. 1.3e-55;  
 Matches 146; Conservative 74; Mismatches 90; Indels 36; Gaps 7;  
 QY 48 LDKLNGRTSGVDARTRPNEKGPVNVNTCNIEFSSSVTKTMDYRVNVLFRQNMNDPL 107  
 DB 37 IDSLT---KGYDIRLRPFCGAPLEIGLEVLASDSISEVMDTITMYLQNYRDEL 93  
 QY 108 STREPPDSDLDP-----SMDLTKPDLFRANEKANFHEVTTDNKLIIRFK 156  
 DB 94 QF--IFNSLSDGENRSVTTMTLTGAFKAIWPTFLANDRNSFLHDTTEKKNKVRILG 151  
 QY 157 NGNVYSYRLTLILSCLDLKNFPMIDIOCTCMQLSSSLICPLSLSLSVGYTKKDLVF 216  
 DB 152 NSGLVYGRFTTTLKCMDLNHPDHOCTVEIE-----SYGYTMDIVL 197  
 QY 217 EWLEDPANQVAEGTLTLPQFILRDEKDLCCCTKHYNTGKTCIEVKFHLEROMGYVLLQ 276  
 DB 198 VYLNDRGAVTGEVDSLPQFSITNVATIN-KIEELSTGQYQLSLIFQORNIIGYFIPT 256  
 QY 277 YIPSLIYILSVSWINNDAPARVAGCITVLTMTTOSSSRSRLPVSIVYKADITWM 336  
 DB 257 YLPSTLIYVLSWISFMINHEATSRVAVLGIITVLTMTTISNIVRSIPRISYVKAIDIVL 316

QY 337 AVCLFEVFAALLEVAINEFV---SHQKFEIRLROR--RORLEE 377  
 DB 317 VACFEVFAALLEVAIVNTTGWGAKRKRKRKRKRKRKRKRKRKR 362  
 RESULT 15  
 ID GABA\_CHICK STANDARD; PRT; 488 AA.  
 AC P24045;  
 DT 01-MAR-1992 (Rel. 21, Created)  
 DT 01-MAR-1992 (Rel. 21, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Gamma-aminobutyric-acid receptor beta-4 subunit precursor (GABA(A)  
 DE receptor).  
 GN GABRB4.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE-91162222; PubMed-1848280;  
 RA Bateson A.N., Lasham A., Darlison M.G.;  
 RT "Gamma-aminobutyric acid A receptor heterogeneity is increased by  
 alternative splicing of a novel beta-subunit gene transcript.";  
 RL J. Neurochem. 56:1437-1440(1991).  
 CC -1- FUNCTION: GABA, THE MAJOR INHIBITORY NEUROTRANSMITTER IN THE  
 CC VERTEBRATE BRAIN, MEDIATES NEURONAL INHIBITION BY BINDING TO THE  
 CC GABA/BENZODIAZEPINE RECEPTOR AND OPENING AN INTEGRAL CHLORIDE  
 CC CHANNEL.  
 CC -1- SUBUNIT: GENERALLY PENTAMERIC. THERE ARE FIVE TYPES OF GABA(A)  
 CC RECEPTOR CHAINS: ALPHA, BETA, GAMMA, DELTA, AND RHO.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: BETA-4 AND BETA-4' (SHOWN HERE);  
 CC ARE PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: X56647; CAA39669.1; -  
 CC EMBL: X56648; CAA39970.1; ALT\_SEQ.  
 DR PIR: JH0360; JH0360.  
 DR PIR: JH0359; JH0359.  
 DR InterPro: IPR000188; GABA<sub>A</sub>receptor.  
 DR InterPro: IPR001175; Neur\_chan.  
 DR Pfam: PF02931; Neur\_chan\_LBD; 1.  
 DR Pfam: PF02932; Neur\_chan\_memb; 1.  
 DR PRINTS: PR00252; NRIONCHANNEL.  
 DR TIGRFS: TIGR00860; LIC; 1.  
 DR PROSITE: PS00236; NEUROR\_ION\_CHANNEL; 1.  
 DR Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;  
 KM Multigene family; Transmembrane; Alternative splicing.  
 FT SIGNAL 1 25  
 FT CHAIN 26 488  
 FT DOMAIN 26 244  
 FT TRANSMEM 245 266  
 FT TRANSMEM 271 292  
 FT TRANSMEM 304 326  
 FT DOMAIN 327 465  
 FT TRANSMEM 466 487  
 FT CARBOHYD 32 32  
 FT CARBOHYD 104 104  
 FT CARBOHYD 173 173  
 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 N-LINKED (GLCNAC. . .) (POTENTIAL).



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: June 25, 2003, 17:13:30 : Search time 52.0671 Seconds  
(without alignments)  
1705.612 Million cell updates/sec

Title: US-10-075-846-4  
Perfect score: 2252  
Sequence: 1 MTLVPAFLSLLMLTLPQ.....POPPAPLREGETTRKLYVD 431

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues  
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

## Database :

SPREMBL\_21:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: -sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_protist:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rviro:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query length	DB ID	Description
1	1886	83.7	416 11	08VHF3 mus musculus
2	1749.5	77.7	456 13	08VHF3 mus musculus
3	1598	71.0	480 11	09J9C9 ratius norv
4	1578	70.1	463 11	09J9C9 ratius norv
5	1570	69.7	449 6	09GKFO mus musculus
6	1570	69.7	449 6	09GKFO mus musculus
7	1555.5	69.1	459 13	08OC93
8	1537.5	68.3	450 13	08OC93
9	1381	61.3	426 13	09OWT3
10	876.5	38.9	498 13	09OWT3
11	850.5	37.8	497 6	09J9C9
12	747.5	33.2	486 5	09GKFO
13	743.5	33.0	438 5	046124
14	742.5	33.0	453 5	077295
15	730.5	32.4	456 5	094900
16	729.5	32.4	452 5	076773

17	729	32.4	606 5	09VSV0	09VSV0 drosophila
18	728	32.3	606 5	09BLX8	09BLX8 drosophila
19	724.5	32.2	481 5	018471	018471 heliothis v
20	724.5	32.2	496 5	018468	018468 heliothis v
21	724	32.1	533 5	016896	016896 aedes aegypt
22	722.5	32.1	456 5	09G053	09G053 drosophila
23	722	32.0	430 5	095097	095097 caenorhabd
24	721	32.0	430 5	017369	017369 caenorhabd
25	720	32.0	601 5	092138	092138 drosophila
26	719.5	31.9	454 5	09VDU5	09VDU5 drosophila
27	718	31.9	421 5	046123	046123 haemochus
28	717	31.8	454 5	0961R4	0961R4 drosophila
29	716	31.8	430 13	090590	090590 gallus gall
30	713	31.7	541 5	09U9B8	09U9B8 ceratilis c
31	712.5	31.6	550 5	017145	017145 lucilia cup
32	711.5	31.6	494 5	09X2M0	09X2M0 heliothis v
33	710.5	31.5	416 5	095096	095096 caenorhabd
34	708.5	31.5	416 5	017367	017367 caenorhabd
35	708	31.4	336 5	09G052	09G052 drosophila
36	706.5	31.4	351 5	025634	025634 onchocerca
37	705.5	31.3	1106 5	022637	022637 caenorhabd
38	704	31.3	379 5	096964	096964 ascaris suu
39	700.5	31.1	478 5	095P06	095P06 caenorhabd
40	700.5	31.1	657 5	09TW41	09TW41 caenorhabd
41	700	31.1	519 13	09DD09	09DD09 brachydanio
42	695.5	30.9	478 5	017548	017548 caenorhabd
43	695.5	30.9	657 5	017547	017547 caenorhabd
44	690.5	30.7	435 5	09TY66	09TY66 haemochus
45	690	30.6	484 5	09U990	09U990 caenorhabd

## ALIGNMENTS

## RESULT 1

08VHF3 PRELIMINARY: PRT: 416 AA.  
ID 08VHF3  
AC 08VHF3  
DT 01-MAR-2002 (TREMBLrel. 20. Created)  
DT 01-MAR-2002 (TREMBLrel. 20. Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21. Last annotation update)  
DE Glycine receptor alpha 4 subunit.  
GN GLRA4.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=B6C3/FE; TISSUE=SPINAL CORD;  
RA Groemer T.-W.W., Becker C.-M.M., Becker K.;  
RT "Localization of different glycine receptor isoforms in murine spinal cord."  
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF462147; AAL69899.1; -  
DR InterPro: IPR000188; GABAA\_receptor.  
DR InterPro: IPR001175; Neur\_chan.  
DR Pfam: PF02931; Neur\_chan\_LBD; 1.  
DR PRINTS: PR00252; Neur\_chan\_memb; 1.  
DR TIGRFAMs: TIGR00860; LIC; 1.  
DR PROSITE: PS00236; NEUROTR\_ION\_CHANNEL; UNKNOWN\_1.  
KW Receptor.  
SO SEQUENCE  
SEQUENCE 416 AA: 48322 MW: C095998PAC077451 CRC64:  
Query Match 83.7%; Score 1886; DB 11; Length 416;  
Best Local Similarity 91.8%; Pred. No. 3.7e-169;  
Matches 358; Conservative 10; Mismatches 8; Indels 14; Gaps 1;

QY 42 MSPSFLDKMGRRISGVARIRPNKGPVAVTCHIFINSSSYTKTMDYRVAVFLRQ 101  
DB 1 MSPSFLDKMGRRISGVARIRPNKGPVAVTCHIFINSSSYTKTMDYRVAVFLRQ 60



```
Db 144 KPDLFANEGKANEHEVTTDKKLIRFKNGNVLSIRLTILSCPMDLKNPMDVQTCIM 203
Oy 189 QLESSLILCSPLSLSVGTMDLVFEMLEDAPAVOVAGLTLPOFLILDEKDLGCT 248
Db 204 QLE-----SFGYTMNDLIFEWODEAP-VQVAGLTLPOFLILKEEKDLRCT 248
Oy 249 KHYTGKFTCIENVFHLEROMGYLLIOMYIPSLILVLSWVSPFINDDAPARALGITT 308
Db 249 KHYTGKFTCIENVFHLEROMGYLLIOMYIPSLILVLSWVSPFINDDAPARALGITT 308
Oy 309 VLTMTTOSGSRASLPKVSYYKAIIDIMAVCLLFFVALLLEYAANFVSRQHKFFIRLR 368
Db 309 VLTMTTOSGSRASLPKVSYYKAIIDIMAVCLLFFVALLLEYAANFVSRQHKFFIRLR 368
Oy 369 RQRQ-----RLE-----EDIIQESRFYRGGLGHCLQARG-----GPMESGITYSP 412
Db 369 KRKKKTAFALEKFRFSDDDEVRESRFSFTAYGMGPCIQAKGVVPGPNHAYV--- 425
Oy 413 QPPAPLREGFTTRKLYVD 431
Db 426 -----MKRSADMKKVFID 439
```

## RESULT 4

```
091XP5 PRELIMINARY; PRT: 463 AA.
AC 091XP5;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Glycine receptor alpha 3 subunit...
GN GLRA3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=SPINAL CORD;
RA Noegel S., Becker C., Becker K.;
RT "Different glycine receptor isoforms are expressed in murine
RT cerebellum."
RL Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF362764; AAK51962.1; -.
DR MGD: MGI:95749; Glra3.
DR InterPro: IPR000188; GABAA_receptor.
DR InterPro: IPR001175; Neur_channel.
DR Pfam: PF02931; Neur_chan_LBD; 1.
DR Pfam: PF02932; Neur_chan_memb; 1.
DR PRINTS: PRO1620; GABAARGAMMA.
DR TIGRFAMs: TIGR00860; LIC; 1.
DR PROSITE: PS00236; NEUROTR_ION_CHANNEL; UNKNOWN.1.
KW Receptor.
SQ SEQUENCE 463 AA: 53605 MW: 889D67DE9B223605 CRC64;
```

Query Match 70.1%; Score 1578; DB 11; Length 463;  
Best Local Similarity 73.0%; Pred. No. 4.5e-140;  
Matches 311; Conservative 33; Mismatches 42; Indels 40; Gaps 8;

```
Oy 22 LLRLVALAKEEYKSGTSGSPSPDFDLKMGRTSGVDARIRPNFKGPVAVTNCINIFNS 81
Db 21 LLSLVATKKTETNARSRSAPSPDFDLKMGRTSGVDARIRPNFKGPVAVTNCINIFNS 80
Oy 82 FSSVTKTTMDYRVAVFLRQONNDRLSLREYPPDSDLDPMSLDSIKKPDILFFANEGKAN 141
Db 81 FGSIAETTMADYRVAVFLRQONNDRLSLREYPPDSDLDPMSLDSIKKPDILFFANEGKAN 139
Oy 142 FHEVTTDKKLIRFKNGNVLSIRLTILSCPMDLKNPMDVQTCIMQLESSLILCSPL 201
Db 140 FHEVTTDKKLIRFKNGNVLSIRLTILSCPMDLKNPMDVQTCIMQLESSLILCSPL 189
Oy 202 SLISLVGTMDLVFEMLEDAPAVOVAGLTLPOFLILDEKDLGCTKHYTGKFTCIENV 261
```

```
Db 190 ----SFGYTMNDLIFEWODEAP-VQVAGLTLPOFLILKEEKDLRCTKHYTGKFTCIENV 244
Oy 262 KFHLEKMGYLLIOMYIPSLILVLSWVSPFINDDAPARALGITTVLTMTTOSGSR 321
Db 245 RFHLEKMGYLLIOMYIPSLILVLSWVSPFINDDAPARALGITTDLTMTTOSGSR 304
Oy 322 SLPKVSYYKAIIDIMAVCLLFFVALLLEYAANFVSRQHKFFIRLRORRO-----RLE- 376
Db 305 SLPKVSYYKAIIDIMAVCLLFFVALLLEYAANFVSRQHKFFIRLRORRO-----RLE- 364
Oy 377 -----EDIIQESRFYRGGLGHCLQARG-----GPMESGITYSPQPPAPLREGFTT 425
Db 365 FYRSPDDEVRESRFSFTAYGMGPCIQAKGVVPGPNHAYV---MKRSP-----DEM 416
Oy 426 RKLKLYVD 431
Db 417 RKVFID 422
```

## RESULT 5

```
09GKFO PRELIMINARY; PRT: 449 AA.
AC 09GKFO;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Glycine receptor alpha 1 subunit Isoform b.
GN GLRA1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21109390; PubMed=11178872;
RA Pierce K.D., Handford C.A., Morris R., Vafa B., Dennis J.A.,
RA Healy P.J., Schofield P.R.;
RT "A nonsense mutation in the alpha subunit of the inhibitory glycine
RT receptor associated with bovine myoclonus."
RL Mol. Cell. Neurosci. 17:354-363(2001).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.
DR EMBL: AF268366; AAG41140.1; -.
DR EMBL: AF268358; AAG41140.1; JOINED.
DR EMBL: AF268359; AAG41140.1; JOINED.
DR EMBL: AF268360; AAG41140.1; JOINED.
DR EMBL: AF268361; AAG41140.1; JOINED.
DR EMBL: AF268362; AAG41140.1; JOINED.
DR EMBL: AF268363; AAG41140.1; JOINED.
DR EMBL: AF268364; AAG41140.1; JOINED.
DR EMBL: AF268365; AAG41140.1; JOINED.
DR EMBL: AF268366; AAG41140.1; JOINED.
DR InterPro: IPR000188; GABAA_receptor.
DR InterPro: IPR001175; Neur_channel.
DR Pfam: PF02931; Neur_chan_LBD; 1.
DR Pfam: PF02932; Neur_chan_memb; 1.
DR PRINTS: PRO1620; GABAARGAMMA.
DR PRINTS: PRO0252; NRIONCHANNEL.
DR TIGRFAMs: TIGR00860; LIC; 1.
DR PROSITE: PS00236; NEUROTR_ION_CHANNEL; 1.
KW Glycopolymers; Ionic channel; Postsynaptic membrane; Receptor;
KW Transmembrane.
SQ SEQUENCE 449 AA: 51676 MW: DC75FB56F3114A4C CRC64;
```

Query Match 69.7%; Score 1570; DB 6; Length 449;  
Best Local Similarity 72.6%; Pred. No. 2.5e-139;  
Matches 307; Conservative 37; Mismatches 53; Indels 26; Gaps 8;

```
Oy 12 LLMTLTPROQVLLRLVALAKEEYKSGTSGSPSPDFDLKMGRTSGVDARIRPNFKGPV 71
Db 9 LYLWE--TIVFSLASKAEAAARSASKPMSPSDFDLKMGRTSGVDARIRPNFKGPV 65
```



Query Match	68.3%	Score 1537.5	DB 13	Length 450
Best Local Similarity	72.9%	Pred. No. 2.9e-136		
Matches 301; Conservative	33	Mismatches 54	Indels 25	Gaps 5

RESULT 9			
ID	ORGWMT3	PRELIMINARY:	PRT: 426 AA.
AC	090WT3;		
DT	01-DEC-2001 (TREMBLrel. 19, Created)		
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)		
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)		
DE	Glucagon receptor alpha24 subunit (Fragment).		
GN	GLRA4B OR GLYRALPHA4.		
OS	Brachydanio rerio (Zebrafish) (zebra danio).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;		
OX	Cyprinidae; Danio.		
RN	NCBI_TaxID=7955;		
RP	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=BRAIN;		
RX	IMBoden M., Devignot V., Goblet C.;		
RA	Medline=21541038; PubMed=11685575;		
RT	"Phylogenetic relationships and chromosomal location of five distinct		
RT	glucagon receptor subunit genes in the teleost danio rerio.";		
RL	Dev. Genes Evol. 211:415-422(2001).		
DR	EMBL; AJ308517; CAC38838.1;		
DR	ZFIN; ZDB-GENE-020402-2; glra4b.		
DR	InterPro; IPR000118; GABA_A_receptor.		
DR	InterPro; IPR001175; Neur_channel.		
DR	Pfam; PF02931; Neur_chan_LBD. 1.		
DR	Pfam; PF02932; Neur_chan_memb. 1.		
DR	TIGRFAMS; TIGR00860; LIC. 1		
KW	PROSITE; PS00236; NEUROTR_ION_CHANNEL; UNKNOWN_1.		
KW	Receptor.		
FT	NON_TER 1		
FT	SEQUENCE 426 AA; 48503 MW; 9187E20B8793234B CRC64;		

Query Match	61.3%	Score 1881	DB 123	Length 426
Best Local Similarity	67.7%	Pred No. 1,5e-123		
Matches 273	Conservative 31	Mismatches 41	Indels 58	Gaps 5
QY	68	GPPAVNTCNIETNSSVTKTMDYRVNVFLRQOMNDPLSLYREYPPDSDLDDPSMDSI	127	
Db	1	GPPAVNTCNIETNSSGITETMDRYLVNLFLOOMNDPLSLAKYEPDDSLDDPSMDSI	60	
QY	128	WRPDLFFANKEGKANGHEVTTDKLRIEFKGNVLVSIRTLTLSCMLDKNFPMIDTOST	187	
Db	61	WRPDLFFANKEGKANGHEVTTDKLRIEPGNVLVSIRTLTLSCMDLKNFPMIDTOST	120	
QY	168	MOLESSILCSPLPSLSVGTMTMDVYEMNLEDAPAVOAGGLTLPQITLIDENDLGGC	247	
Db	121	VOLE-----SFGYTMDDLFLQWIDEGP-VQVADDLMIPOCVLKEEKDLGIC	165	
QY	248	TKHYNTGKETIEVAFHLEROMGYLLIOMYISLVLIVLSWVSFWINMDAARVGLGIT	307	



Db 166 ARHNTGKFTGCIIEVKEFLEROMAYVLIOMVIPSLLTVLISVSWFWMINDAARVGLGIT 225  
QY 308 TVLMTTQSSGSRASLPKVSVAKIDIMAVCLLEFVALLLEYAAINFSROHKEFRLR 367  
Db 226 TVLMTTQSSGSRASLPKVSVAKIDIMAVCLLEFVALLLEYAAINFSROHKEFRLR 285  
QY 368 RRRORR-----RLEEDIIQESFFPR 388  
Db 286 RRLROEORNRASGQGAETKNSNNVTGTPSRNAORQCSACAREEOLASQNDLPP 345  
QY 389 GYGLCHICLARDGGMEGSGIYSPOPALLREGETTRLYND 431  
Db 346 GFGIDTSLSG-DGPLSEAAAFRGLPPHALFD--IRREVE 384

## RESULT 10

Q9DES8 PRELIMINARY: PRT: 498 AA.

AC 09DES8: 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)  
DE 01-JUN-2002 (TREMBLrel. 21, last annotation update)  
GN Glycine receptor beta2 subunit precursor (Fragment).  
OS Brachydanio rerio (Zebrafish) (Zebra danio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_Taxid=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=BRAIN.  
RA Imboden M., Devignot V., Korn H., Goblet C.;  
RT Regional Distribution of Glycine Receptor mRNA in the Central Nervous  
System of Zebrafish.";  
Submitted (Jul-2000) to the EMBL/GenBank/DBJ databases.  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (By similarity).  
EMBL: AJ404971; CCNC1688.1; -  
DR ZFIN: ZDB-GENE-010410-2; g1rb.  
DR InterPro: IPR000188; GABA\_A\_receptor.  
DR InterPro: IPR001175; Neur\_channel.  
DR Pfam: PF02932; Neur\_chan\_LBD; 1.  
DR Pfam: PF02932; Neur\_chan\_mem; 1.  
DR PRINTS: PRO1620; GABARGAMA.  
DR PRINTS: PRO0252; NRIONCHANNEL.  
DR TIGRFAMs: TIGR00860; LIC; 1.  
DR PROSITE: PS00236; NEURORF\_ION\_CHANNEL; 1.  
KW Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor; Signal;  
KW Transmembrane.  
FT NON\_TER 1  
FT SIGNAL <1 22 POTENTIAL.  
SQ SEQUENCE 498 AA; 56358 MW; 12436FC344F0BB CRC64;

Query Match 38.9%; Score 876.5; DB 13; Length 498;  
Best Local Similarity 49.1%; Pred. No. 6, 4e-74;

Matches 189; Conservative 59; Mismatches 96; Indels 41; Gaps 11;

QY 17 LFGQVLLR--VALAKEE--VKSQTGSGSPMSPDF-----LDKIMRTSG-----Y 58  
Db 8 LKGLILMLLVQFSAEGRKPKKKGKQYICPSQASADLDVRVANSSTNINRLMKTY 67  
QY 59 DARIRPNFKGPPVNTCNIFINSFSSVTKTMDYRVNVELRQOOWNDPRLSY-REYPDSTL 117  
Db 68 DSRIIRNFKGIEVEDKVINIFINSFSGIOETMDYRVNIFLRQWMDPRLRLPDRKSDL 127  
QY 118 DDDPSLDSIKRPDLFEFANEKANHEVTDDKLLRIFKNGVNLISIRLLTSLCLMDIK 177  
Db 128 TVDPKPFQCLMKRPDLFEFANEKANHEVDYQENILFIFRNGVNLISMRSLVTLSCPLAQ 187  
QY 178 NFPPMOTCTMOLESSIICSPPLSLSGVYTMQDLVEWLEDAVAVOAGLTLPPQI 237  
Db 188 LFPMDIOTCKMDE-----SFGITTKDLVEMKOSGDP-VOMDE-TALPQD 231

QY 238 LRDEK-DLGCCTKHY-NTGKFTGCIIEVKEFLEROMGYLLIOMVIPSLLIYLSVSWFWM 295  
Db 232 VQEOEDKYNANCKRFTYGTCTYTCVEYIFTLRQOVGMGYAFTLLIYLSVLSWINP 291  
QY 296 DAAPAVGLGITVTVLMTTQSSGSRASLPKVSVAKIDIMAVCLLEFVALLLEYAAI-- 353  
Db 292 DASARVPLGILSVLTSSECTSLASELPKVSVAKIDIMAVCLLGFASIVEYAVOV 351  
QY 354 --NFSROHKEFRLRRORRQLE 376  
Db 352 MNSPKRIEAEKIKMAKEKAREKE 376

## RESULT 11

Q9GJS9 PRELIMINARY: PRT: 497 AA.

AC 09GJS9: 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)  
DE 01-JUN-2002 (TREMBLrel. 21, last annotation update)  
GN Glycine receptor beta subunit.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_Taxid=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Pierce K.D., Handford C.A., Morris R., Vafa B., Dennis J.A.,  
RT Healy P.J., Schofield P.R.;  
RT "Bovine Myoclonus is Caused by a Mutation in the  $\alpha 1$  Subunit of the  
Inhibitory Glycine Receptor.";  
Submitted (May-2000) to the EMBL/GenBank/DBJ databases.  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (By similarity).  
EMBL: AF268376; AAG14347.1; -  
DR EMBL: AF268374; AAG14345.1; -  
DR EMBL: AF268367; AAG14345.1; JOINED.  
DR EMBL: AF268368; AAG14345.1; JOINED.  
DR EMBL: AF268369; AAG14345.1; JOINED.  
DR EMBL: AF268370; AAG14345.1; JOINED.  
DR EMBL: AF268371; AAG14345.1; JOINED.  
DR EMBL: AF268372; AAG14345.1; JOINED.  
DR EMBL: AF268373; AAG14345.1; JOINED.  
DR InterPro: IPR000188; GABA\_A\_receptor.  
DR InterPro: IPR001175; Neur\_channel.  
DR Pfam: PF02931; Neur\_chan\_LBD; 1.  
DR Pfam: PF02932; Neur\_chan\_mem; 1.  
DR PRINTS: PRO1620; GABARGAMA.  
DR PRINTS: PRO0252; NRIONCHANNEL.  
DR TIGRFAMs: TIGR00860; LIC; 1.  
DR PROSITE: PS00236; NEURORF\_ION\_CHANNEL; 1.  
KW Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor;  
KW Transmembrane.  
SQ SEQUENCE 497 AA; 56039 MW; 82F140C115A887E6 CRC64;

Query Match 37.8%; Score 850.5; DB 6; Length 497;  
Best Local Similarity 47.1%; Pred. No. 1, 8e-71;

Matches 184; Conservative 65; Mismatches 89; Indels 53; Gaps 11;

QY 9 LSFLLMTLPGQVLLVALAKEEVSGTGSGSPM-----SPSPFLDK 50  
Db 7 VAFELIISL-----CLIEAIVSKSSKKGKQYICPSQASADLDARVANSSTNINRL 62  
QY 51 LKGRISGYDARIRPNFKGPPVNTCNIFINSFSSVTKTMDYRVNVELRQOOWNDPRL--- 107  
Db 63 LL--VSYPRIIRPNFKGIPVDVYVNIIFINSFSGIOETMDYRVNIFLRQWMDPRLKLP 119  
QY 108 -SYRETPDSDLDLPSMDLSIKRPDLFEFANEKANHEVTDDKLLRIFKNGVNLISIRL 166  
Db 120 SDFR--GSDALTVDPTRYKLMKRPDLFEFANEKANHEVDYQENILFIFRQDDVIVSMRL 177





DB 295 SLGVTLLTMTQASINSKLPVSYIKADVIGVCLAFIGALLLEYAVVNYGR--KE 352  
 OY 363 FTRLRRORORREEDIQESR 384  
 DB 353 FLR-KEKKKTRLDCCPSER 373

## RESULT 14

077295 PRELIMINARY; PRT; 453 AA.  
 ID 077295;  
 AC 077295;  
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE DROSGLUCL.  
 GN GLUCL-ALPHA OR INDEFINITE OR CG7535.  
 OS Drosophila melanogaster (fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NCBI\_Taxid=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Semenov E.P., Pak W.L.;  
 RT "Diversification of Drosophila Chloride Channel Gene by multiple  
 posttranscriptional mRNA modifications";  
 RL J. Neurochem. 0:0-0(0).  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.  
 DR EMBL: AA002232; CA053260.1; -;  
 DR Flybase: FBgn0024963; GluCl-alpha.  
 DR InterPro: IPR001188; GABA\_receptor.  
 DR InterPro: IPR001175; GABA\_receptor.  
 DR Pfam: PF02931; Neur\_chan\_LBD; 1.  
 DR Pfam: PF02932; Neur\_chan\_mem; 1.  
 DR PRINTS: PR00252; NRIONCHANNEL.  
 DR TIGRfams: TIGR00860; LIC; 1.  
 DR PROSITE: PS00236; NEURORF\_ION\_CHANNEL; 1.  
 DR GlycoProtex: Ionic channel; Postsynaptic membrane; Transmembrane.  
 KW KMWQ  
 SQ SEQUENCE 453 AA; 51905 MW; C0AD53FA6968C2A CRC64;

Query Match 33.0%; Score 742.5; DB 5; Length 453;  
 Best local Similarity 43.1%; Pred. No. 2.4e-61;  
 Matches 157; Conservative 58; Mismatches 120; Indels 29; Gaps 7;

OY 21 VLLRVALAKEEYKSGTSGSPSPDLDLKGRTSGYDARIP--NFKGPPVNTCNI 77  
 DB 9 ILTFASLCSASLANNAKVNFREREKKVLDQILG-AGKYDARIRPSGINGTDGPAVVYNI 67  
 OY 78 FINSFSSVTKTMDYRVNVLQOQNDPRLSYREYPPDSDLDLPPSLDSTINKPDLFPANE 137  
 DB 68 FVRSIKSIDVTMEYSVQLTFREQWTDERLKFDDIGRLKYLTLTFANVWMPDLFFSNE 127  
 OY 138 KGANFHEVTTDKLIRFKNGVLSIRLTLLSCMLDKNPPMDIQCTMOLESSILC 197  
 DB 128 KEGHFNIITIMPVYIRIFPNSGLYSIRISLTACPMNKLYPLDRQICSLM----- 180  
 OY 198 SPLPSTLSISVGYTMKDLVEFWELEDAVAVOAGLTLPOFLRDEKDL-GCCTRKHYNTGKF 256  
 DB 181 -----ASYGWTNDLVLFMKEGDP-VGVVKNLHLPRFTL--EKLFDYCSKNTNGEY 230  
 OY 257 TCIIEVGFHLEROMGYLLIOMYIPSLILVILSVSWFMINNDAPARVGLGITVLTWTQS 316  
 DB 231 SCLKVDLFRRESYLLIYIPCCMLVIVSWVSFPLDGAIPARVGLGITVLTWTQS 290  
 OY 317 SGRASLSPKVSIVKADIMAVCLLFEVALLLEYAIVNSR-----QKKEFTLRRLR 369  
 DB 291 SGINSLPVPVSTKADIVMTGCLTFVGFALLEFALVYNASRSGSKNANMKHKKRR 350  
 OY 370 QRRQ 373  
 DB 351 DLEQ 354

## RESULT 15

094900 PRELIMINARY; PRT; 456 AA.  
 ID 094900;  
 AC 094900;  
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Glutamate-gated chloride channel.  
 GN GLUCL-ALPHA OR DROSGLUCL OR CG7535.  
 OS Drosophila melanogaster (fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NCBI\_Taxid=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN-OREGON R; TISSUE-HEAD;  
 RC MEDLINE-96355334; PubMed-8702744;  
 RX Cullis D.F., Pares P.S., Liu K.K., Schaeffer J.M., Arena J.P.;  
 RT "Identification of a Drosophila melanogaster glutamate-gated chloride  
 channel sensitive to the antiparasitic agent avermectin";  
 RL J. Biol. Chem. 271:20187-20191(1996).  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.  
 DR EMBL: U58776; AAC47266.1; -;  
 DR Flybase: FBgn0024963; GluCl-alpha.  
 DR InterPro: IPR001188; GABA\_receptor.  
 DR InterPro: IPR001175; Neur\_chan\_LBD; 1.  
 DR Pfam: PF02931; Neur\_chan\_LBD; 1.  
 DR Pfam: PF02932; Neur\_chan\_mem; 1.  
 DR PRINTS: PR00252; NRIONCHANNEL.  
 DR TIGRfams: TIGR00860; LIC; 1.  
 DR PROSITE: PS00236; NEURORF\_ION\_CHANNEL; 1.  
 DR GlycoProtex: Ionic channel; Postsynaptic membrane; Transmembrane.  
 KW KMWQ  
 SQ SEQUENCE 456 AA; 52347 MW; 4008E5765756B1C CRC64;

Query Match 32.4%; Score 730.5; DB 5; Length 456;  
 Best local Similarity 42.3%; Pred. No. 3.2e-60;  
 Matches 154; Conservative 61; Mismatches 120; Indels 29; Gaps 7;

OY 21 VLLRVALAKEEYKSGTSGSPSPDLDLKGRTSGYDARIP--NFKGPPVNTCNI 77  
 DB 10 ILTFASLCSASLANNAKVNFREREKKVLDQILG-AGKYDARIRPSGINGTDGPAVVYNI 68  
 OY 78 FINSFSSVTKTMDYRVNVLQOQNDPRLSYREYPPDSDLDLPPSLDSTINKPDLFPANE 137  
 DB 69 FVRSIKSIDVTMEYSVQLTFREQWTDERLKFDDIGRLKYLTLTFANVWMPDLFFSNE 128  
 OY 138 KGANFHEVTTDKLIRFKNGVLSIRLTLLSCMLDKNPPMDIQCTMOLESSILC 197  
 DB 129 KEGHFNIITIMPVYIRIFPNSGLYSIRISLTACPMNKLYPLDRQICSLM----- 181  
 OY 198 SPLPSTLSISVGYTMKDLVEFWELEDAVAVOAGLTLPOFLRDEKDL-GCCTRKHYNTGKF 256  
 DB 182 -----ASYGWTNDLVLFMKEGDP-VGVVKNLHLPRFTL--EKLFDYCSKNTNGEY 231  
 OY 257 TCIIEVGFHLEROMGYLLIOMYIPSLILVILSVSWFMINNDAPARVGLGITVLTWTQS 316  
 DB 232 SCLKVDLFRRESYLLIYIPCCMLVIVSWVSFPLDGAIPARVGLGITVLTWTQS 291  
 OY 317 SGRASLSPKVSIVKADIMAVCLLFEVALLLEYAIVNSR-----QKKEFTLRRLR 369  
 DB 292 SGINSLPVPVSTKADIVMTGCLTFVGFALLEFALVYNASRSGSKNANMKHKKRR 351  
 OY 370 QRRQ 373  
 DB 352 DLEQ 355

Search completed: June 25, 2003, 17:16:58  
 Job time : 54.0671 secs

---

**THIS PAGE BLANK (USPTO)**

Db	510	CGATATCCCTTCAGATTTTCTTGACAAAGTTAATGGGAAGACATCGAGTATGATGCAAG	569
Oy	183	GATTCGGCCCAATTTTAAAGGCCACCGGTGAACGTGACTGCAACATCTTCATCAACG	242
Db	570	AATCAGGCCAAATTTTAAAGAGTCTCTCAATAACGTTACTTCCAAATATTTTATCAACG	629
Oy	243	TTTTCAGCTCCGTACCAACGACCAATGGAATGCACTACCGGGTGAATGTCCTCTTGCCGCAACA	302
Db	630	TTTTGGATCTACGATCAGAAACGACATGCGTACCCAGAGAAATATTTTCTGAGCAACAA	689
Oy	303	GTGGAATGACCCACAGCGCTGTCCATCCGAGAAATATCCGTATGACTCTCTGGACCTGATNC	362
Db	690	GTGGAATGATTTACGCGGTGGCGTACAGTGAATGATACCCAGATGACTCCCTGGACTTGGACCC	749
Oy	363	CTCATGCTGGACTCTATCTATGGAAGCCAGACCTTTCTTGTCTAATGAGAAAGGGCCAA	422
Db	750	ATCCATGCTAGACCTCCATTTGGAAACCAAGATTTGTCTTTGCCAATGAGAAAGGGTCCAA	809
Oy	423	CTTCCATGAGGTGACCAACGGAACAAGTTACTGCGCATCTTAAATAATGGGAATGTGCT	482
Db	810	CTTCCACAGATTCACCACTGACCAACAAATGCTAAGGATTTGCAAAAATGGCAAAATGCT	869
Oy	483	GTACAGCATAGGGGTGACCCCTCAATTTGTGCTCGCTGATGAGACTGCAAGAATCTTCCCAT	542
Db	870	CTACAGTATTCAGACTCACTTGACCTTATCTGTGCCATGAGACTTGAAGAATCTTCGAT	929
Oy	543	GGACATCCAGACGTGACAGATGACAGCTTGAGAGCTACATCTGACAGCCCTTGCC	602
Db	930	GGATGTCCAGACCTGTACATGACAGCTGGAGAGT-----	964
Oy	603	ATCTGTCTACTTTCAGTGGCTACACCATGAAAGACCTGCTGTTTGAAGGCGTGAAGA	662
Db	965	-----TTGGGTACAGATGAAGACCTGATATTTGAGTGGTTAAAGTGA	1007
Oy	663	TGCTCTGCTGTCGCAAGTGGCTGAAGGGGCTACTCTGCCCAAGTTTATCTTGGGGATGA	722
Db	1008	TGCTC---CAGTGCAGTGTCTGAAGGATTACCTGCCCAAGTTTATTTGAAGAAGA	1064
Oy	723	GAGGATCTAAGGCTGTGTACCAACACTACACACAGGAATTCACCTGCTCATCAGGT	782
Db	1065	GAGGAACTTGGCTACTCTTACAAAGCACTACACACAGTGAAGTTTACCTGCAATGAGGT	1124
Oy	783	AAAGTTTCACTGGAAGCGGAGATGGGTACTATCTGTATGATGATGATACATCCCAAGCT	842
Db	1125	CAGATTTCACGTGAAGCCCAATGGGATATTATTTGATCCAGATGTCATCCCAAGCT	1184
Oy	843	ACTCATGTCATCTGCTCGGGGTCTCTTGTGATCAACATGATGCTGCGCCCTGCGCG	902
Db	1185	GCTTATAGTAAATTTTGTCTCGGGTTCTCTTTGGATAAATGTGATGAGCCCTGCGCAG	1244
Oy	903	TGTGGGCTGTGGGCATCACCACTGCTACCATGACACCCAGAGCTTGGCTCCCGGC	962
Db	1245	GGTCGACATGGGCATCACCACTGATTAACGATGACACCCAGAGTTTACGGCTCCAGGGC	1304
Oy	963	CTCTTGGCTAAGGTCTCTACGTAAAGGCATGCATCTGGAATGGGTGGTGTGCTCT	1022
Db	1305	ATCTGTCGCAAAAGTCTCTATGTAAAGCGATGTGACATCTTGGATGGCGGTGTGCTTCT	1364
Oy	1023	CTTGTGTTGCGTGCCTTCTGAGAGTATGCTCCATAAATTTTGTCTCGTCAGCATAA	1082
Db	1365	GTTTGTGTTTGTGCTTACTGGAATAGCCAGGGGAACTTCGTCCAGGCAACACAA	1424
Oy	1083	AGAATTCATGCACTTGCAGGAAGACGACAGCGCCCAACGCTTGGAGAGAAATATCATCA	1142
Db	1425	GGAGTGTCTCGCCTCCGAAAGACAGGAAGAGGCGGAATTAAGGAAGAGCTTACTCG	1484
Oy	1143	AGAAAGCTTTTCTATTTTCCGTGGGTATGAGCTTGGGCACTGTCGACGGCAAGAGATGG	1202
Db	1485	TGAAAGTCGTTTAAATTTTAAAGCGGTTATGGATGGGTCACTGCTTCCAAGTGAAGAATGG	1544
Oy	1203	AGGTCCAATGGAAGGTCTGGAATTTATAGTCCCAACCTCCAGGCCCTTTCTAAGGA	1262
Db	1545	AACGCTGTCAAGGCCAC-----ACCTGCAACCGACCTGCCAACCGCCAAAGA	1595

**THIS PAGE BLANK (USPTO)**

QY 1263 AGGAGAAACCGGAACTCTACGTGACTGACGACAGAGATTGACACCATCTCCG 1322  
DB 1596 TGGAGATGCTATCAAGAGAAAGTGTGTGACCGGCAAAAGAGATTGACAGATATCTCG 1655  
QY 1323 GGCCTCTCTCCCTTTCATCTTCCATCTCTCAATATCTCTGAGGTGTCTATTAAGT 1382  
DB 1656 ACCTCCCTTCCCTGAGCTTCTTCTTCAACATCTTCTTCAATGATGATCAACAGAT 1715  
QY 1383 GCTATGCTGACAGATATCCACCAG 1407  
DB 1716 CATTCGGCATGAGATGTCACAG 1740  
RESULT 14  
AF268375 1744 bp mRNA linear MAN 22-MAY-2001  
LOCUS AF268375  
DEFINITION Bos taurus glycine receptor alpha 1 subunit (Glr1) mRNA, complete  
ACCESSION AF268375  
VERSION AF268375.1 GI:10180958  
KEYWORDS  
SOURCE  
ORGANISM Bos taurus.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
Bovidae; Bovinae; Bos.  
1 (bases 1 to 1744)  
Pierce, K.D., Handford, C.A., Morris, R., Vafa, B., Dennis, J.A.,  
Healy, P.J., and Schofield, P.R.  
A nonsense mutation in the alpha subunit of the inhibitory glycine  
receptor associated with bovine myoclonus  
Mol. Cell. Neurosci. 17 (2), 354-363 (2001)  
JOURNAL  
MEDLINE 21109390  
PUBMED 11178872  
2 (bases 1 to 1744)  
Pierce, K.D., Handford, C.A., Morris, R., Vafa, B., Dennis, J.A.,  
Healy, P.J., and Schofield, P.R.  
Direct Submission  
Submitted (17-MAY-2000) Neurobiology Program, Garvan Institute of  
Medical Research, 384 Victoria Street, Sydney, NSW 2010, Australia  
FEATURES  
source  
1..1744  
/organism="Bos taurus"  
/db\_xref="taxon:9913"  
1..1744  
/gene="Glr1"  
229..1602  
/gene="Glr1"  
/note="Glr1ins: alternatively spliced"  
/codon\_start=1  
/product="glycine receptor alpha 1 subunit"  
/protein\_id="AAG14346.1"  
/db\_xref="GI:10180958"  
/translation="MSENTLRLLYMETIVFESLASKAEARASASRPMSPDLK  
LMGRSGYDARLPNFKGPVNVSCNIFNSFSAIETMDNRAVIFLQONDRILA  
YNEYDDSLDPSMDLSIMKDFEFANEGHPEITDNDKLLISRGANVLSIRI  
TLTACPDLDKPNFVOCITCMOLESPFGTMNDLPEQEOGAAVAGITLPOITIK  
BEKDIKYCTKHYNTGKFCIEERHLEKRMGYLLIOMTIPSLIYLISLISWIMDA  
APARVGLITVLYMTOSGSRASLPKYSYKVIDIMWVCLLEFVSLALEYAVNF  
VSRQKELIRFRKRHRKSPMLNLFODEAGEGRNFSAYCMGACLOAKDGISVKG  
ANNSVTTNPPAPSPKSPKEMRLFIORAKKIDKISRIGPPMFLIFNMYIYKIVR  
REDVHNO"  
BASE COUNT 396 a 508 c 451 g 389 t  
ORIGIN  
Query Match 39.4%; Score 646.8; DB 4; Length 1744;  
Best Local Similarity 71.2%; Pred. No. 9, 6e-169;  
Matches 956; Conservative 0; Mismatches 317; Indels 69; Gaps 5;  
DB 299 AGGAGCTGAGCTGCTCGCTCTTCCAGCCCATGTCACCGATTCCTGGATA 358

QY 149 AACTATGAGGCGAATCTGATATGATGCCAGATTGCGCCCAATTTTAAAGGCCAC 208  
DB 359 AACTATGAGGCGAATCTGATATGATGCCAGATTGCGCCCAATTTTAAAGGCCAC 418  
QY 209 CCGTGAAGCTGACCTGCAACATCTTCATCAACAGTTTCAGCTCCGTCAACCAAGCACAA 268  
DB 419 CAGTGAATGTCAGCTCAACATCTTCATCAACAGTTTCAGCTCCGTCAACCAAGCACAA 478  
QY 269 TGGATACCGGCTGAATCTCTCTGCGGCAACAGTGAATGACAGCCCTGCTTAC 328  
DB 479 TGGATACCGGCTGAATCTCTCTGCGGCAACAGTGAATGACAGCCCTGCTTAC 538  
QY 329 GAGATATCTGATGACTCTCTGACCTGATCTCCATGCTGAGCTATCTGGAAGC 388  
DB 539 ATGATATACCGGATGATCTCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGC 598  
QY 389 CAGACCTCTCTCTGCTAATGAGAAAGGCGCAACCTTCATGAGGTGACACGACACAA 448  
DB 599 CTGACCTGCTCTCTGCTAATGAGAAAGGCGCGCAACCTTCATGAGGTGACACGACACAA 658  
QY 449 AGTACTGCGGATCTTCAAGAAATGGGAATGCTGTACAGATCAAGCTGACCTCATTT 508  
DB 659 AACTGCTGAGATCTCCCGGAAAGGGAATGCTGTACAGATCAAGCTGACCTCATTT 718  
QY 509 TGTCTGCGGATGAGACCTGCAAGAACTCCCATGAGATCCAGACGTCAGTCAAGC 568  
DB 719 TGGCTGCGGATGAGACCTGCAAGAACTCCCATGAGATCCAGACGTCAGTCAAGC 778  
QY 569 TTGAGAGCTCATCATCTCTGACACCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCA 628  
DB 779 TGGAGAGCT-----TTGATATCA 796  
QY 629 CCATGAAGACCTGCTGTTTGAAGTGGCTGGAAGATGCTCTCTCTCTCTCTCTCTCTCT 688  
DB 797 CCATGAAGACCTGCTGTTTGAAGTGGCTGGAAGATGCTCTCTCTCTCTCTCTCTCTCT 853  
QY 689 GGCCTGCTGCGGATGAGACCTGCAAGAACTCCCATGAGATCCAGACGTCAGTCAAGC 748  
DB 854 GACTGATCT 913  
QY 749 ACTACACACAGGAAATTCACCTGACGAGTAAAGTTTACCTGACGACGACAGATG 808  
DB 914 ATTACACACAGGAAATTCACCTGACGAGTAAAGTTTACCTGACGACGACAGATG 973  
QY 809 GCTACTATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 868  
DB 974 GCTACTATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1033  
QY 869 CTTTCTGATCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 928  
DB 1034 CTTTCTGATCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1093  
QY 929 TCACATGACACACGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 988  
DB 1094 TCACATGACACACGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1153  
QY 988 AGGCAATGACATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1048  
DB 1154 AGGCAATGACATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1213  
QY 1049 ATGCGCATTAATTTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1108  
DB 1214 ATGCGCATTAATTTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1273  
QY 1109 AGAGCGCC-----AACGCTGAGAGAGATATCATCAAGAAAGTC 1150  
DB 1274 GAAAGATACACAGAGCCCATGTTGATATCTTCCAGAGATGATGAGTGAAGAGCC 1333  
QY 1151 GTTTTATTTTCCGTGCTATGAGCTTGGGCT-----ACTGCTGACGACGAGATGAGGTC 1207  
DB 1334 GCTTCAACTTCTGCGCTTACGACGATGGGCCAGCTGCTTACGACGACGATGATGATG 1393

**THIS PAGE BLANK (USPTO)**



/protein\_id="CAC16687.1"  
/db\_xref="GI:11322384"  
/translation="MTHTSLITINDCINSCRYHVFSGFSSISPLCELYLMOKPMSPS  
DPLDKRMGRSGDARIKPMKGRPVNMTNFINSGSTETTMQRLNRLNVEFROOM  
DPLATSETPDASLDLDPMSLDIMKPDLEFANEKANPHEVTTDKLRIENGVNL  
YSIRLILISCPMDLKNFMDIOTCTMQLESFGYTMNDLLEFVLSNDPVOVADLTLP  
OFVLEKEKEDGYCTKHNTGKTCIEVKFHELEOMGYLLIOMIPSLIYLISWSEW  
INDDAPARVGLITVLTWTOSGSRASLPRYSYKALDIAMAVCLLIVFALLLEY  
AAVNFVROKKEFIRLRKRORIEDIVRESGFEGGLGLOCTODKDAVES  
SVAPPPPOVLYLDEAVKRRFPDRAKRIDTISRATVPPLIFNFVMTTYVLRHE  
DIRANP  
1891. .2045  
/gene="glyr alpha"  
BASE COUNT 524 a 501 c 460 g 560 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 8,72e-168 Length: 2045  
Score: 1749.50 Matches: 334  
Percent Similarity: 84.91% Conservative: 26  
Best Local Similarity: 78.77% Mismatches: 25  
Query Match: 77.69% Indels: 39  
DB: Gaps: 4  
US-10-075-846-4 (1-431) x DREA04970 (1-2045)  
OY 9 LeuSerPheLeuLeuLeuTPTPTPhLeuProLysGlnValLeuLeuArgValAlaLeuAla 28  
DB 610 CTCTGCTTTCTCTATCTCTG----- 630  
OY 29 LysGlnGluValLysSerGlyThrLysGlySerGlnProMetSerProSerAspPheLeu 48  
DB 631 -----CAGAACCTATGTCCTCCATCGACTTCTG 660  
OY 49 AspLysLeuMetGlyArgThrSerGlyThrAspAlaArgLeuProAspPheLysGly 68  
DB 661 GACAGCGATGATGAGCAGAACCTCGGCTATGATCTCCATCAGACCCAACTTCAAGA 720  
OY 69 ProProValAsnValThrCysAsnIlePheIleAsnSerPheSerValThrLysThr 88  
DB 721 CCACACGTAACGTCACCTGTAACATCTTCAACAGATTGGATCTATTACAGACA 780  
OY 89 ThrMetAspTyrArgValAsnValPheLeuArgGlnGlnIlePasnAspProArgLeuSer 108  
DB 781 ACAATGATACAGCGCTAAAGCTCTTCTACGCGCAGCAGTGAACGACCTAGACTGCC 840  
OY 109 TyrArgGluTyrProAspAspSerLeuAspLeuAspProSerMetLeuAspSerIleTyr 128  
DB 841 TACAGTGAATATCCCGATGATCTCTAGACTGACCTTCTATGTTGACTCCATATAG 900  
OY 129 LysProAspLeuPhePheAlaAsnGluLysGlyAlaAsnPheHisGlnValThrThrAsp 148  
DB 901 AAACCTGCTGTTTGTTCCTATATAGAAAGTGGCAACTTCCATGAAGTCAACCAAGC 960  
OY 149 AsnLysLeuLeuArgIlePheLysAsnGlnLysValLeuTyrSerIleArgLeuThrLeu 168  
DB 961 AACAGCTGAGGATCTTTCAGAAATGGAATGTCTTACAGCATCAGGCTACACATC 1020  
OY 169 IleLeuSerCysLeuMetAspLeuLysAsnPheProMetAspIleGlnThrCysThrMet 188  
DB 1021 ATCTGCTGTGTGCTCCATGACTGAGAAATTTCCCAATGAGACATTCAGACTGTACCA 1080  
OY 189 GlnLeuGluSerSerIleLeuCysSerProLeuProSerLeuSerLeuSerValGly 208  
DB 1081 CAGCTAGAA-----AGCTTTGGC 1098  
OY 209 TyrThrMetLysAspLeuValPheGlnTyrLeuGlnAspAlaProAlaValGlnValAla 228  
DB 1099 TACACCATGAACGATCTGATCTTCGAGTGGCTTTCGATTAACCT---GTGCAGGTGGC 1155  
OY 229 GlnGluLeuThrLeuProGlnPheIleLeuArgAspGlnLysAspLeuGlyCysThr 248  
DB 1156 GATGACTGACTCTTCCCTCAGTTTGTACTAAAGGAGAAAGATCTCGGCTACTGCACT 1215

OY 249 LysHisTyrAsnThrGlyLysPheThrCysIleGlnValLysPheHisLeuGlnArgGln 268  
DB 1216 AACCACTACAAACAGAGTAATATCACCTGATGAGGTGAAGTTCACTAGAGACAG 1275  
OY 269 MetGlyTyrTyrLeuIleGlnMetTyrIleProSerLeuLeuIleValIleLeuSerTyr 288  
DB 1276 ATGGCTACATATCTCTTCAAGATGATACATCCAGCTGCTGATTCATTTGTCTTG 1335  
OY 289 ValSerPheThrIleAsnMetAspAlaAlaProAlaArgValGlyLeuGlyIleThrThr 308  
DB 1336 GTGCTTTCTGATACATATGATGATGCGCAGCCGCTGGTGGTATCACCAACC 1395  
OY 309 ValLeuThrMetThrThrGlnSerSerGlySerArgAlaSerLeuProLysValSerTyr 328  
DB 1396 GTGCTGACATGACACCAAGCTGCTGTTCAAGACCTCGTACCAAGTGTCTTAC 1455  
OY 329 ValLysAlaIleAspIleTyrPheMetAlaValCysLeuLeuPheValPheAlaIleLeu 348  
DB 1456 GTCAAGCCCATTTGACATCTGATGCTGTTGCTGCTGATCGTGTGCTGATGTTG 1515  
OY 349 GlnTyrAlaIleAlaAsnPheValSerArgGlnHisLysGlnPheIleArgLeuArg 368  
DB 1516 GAGTATCAGCAGTCACTGCTGCTCAAGCAGCAGCAGGATTCATCAGACTCGCAAG 1575  
OY 369 ArgGlnArgArgGlnArgLeuGlnGlnAspIleIleGlnGlnSerArg---PheTyrPhe 387  
DB 1576 AAGCAGGAGGAGGAGAGATAGAGAGACCTTGAGAGAAACCGGTCTTTACTTC 1635  
OY 388 ArgGlyTyrGlyLeuGlnHisCysLeuGlnAlaArgAspGlyGlyProMetGlnGlySer 407  
DB 1636 CGAGGTACGAGACTGGGCTATGCTTGCMAAAGATGGAGACAGAGTGGAGGGTCC 1695  
OY 408 GlyIleTyrSerProGlnProAlaProLeuLeuArgGlnGlnGlnTyrThrArgLys 427  
DB 1696 AGTATATTCGCCCAACCACTTCAGTCAAGTGTTCACGAGAGAGCCGTCGGAG 1755  
OY 428 LeuTyrValAsp 431  
DB 1756 CGCTTCGTGAC 1767  
RESULT 5  
LOCUS HSGLYRAL  
DEFINITION "HSGLYRAL alpha-2 strychnine binding subunit of inhibitory glycine receptor mRNA."  
ACCESSION X52008.1 GI:31848  
VERSION X52008.1  
KEYWORDS glycine receptor; inhibitory glycine receptor; strychnine binding.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens.  
REFERENCE 1 (bases 1 to 1857)  
AUTHORS Greeningloh,G., Schmieden,V., Schofield,P.R., Seeburg,P.H., Sidique,T., Mohandas,T.K., Becker,C.M. and Betz,H.  
TITLE Alpha subunit variants of the human glycine receptor: primary structures, functional expression and chromosomal localization of the corresponding genes  
JOURNAL EMBO J 9 (3), 771-776 (1990)  
MEDLINE 90183975  
PUBMED 2155780  
FEATURES  
source location/Qualifiers  
1..1857  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/tissue\_type="brain"  
/dev\_stage="fetal"  
388..1746  
/gene="GLYRAL"  
388..1746  
/note="strychnine binding alpha-2 subunit".

THIS PAGE BLANK (USPTO)

```
/codon-start-1
/product="Inhibitory glycine receptor"
/protein_id="CAA36257.1"
/db_xref="GI:31849"
/db_xref="SWISS-PROT:P23416"
/translation="MKROLNLTALFAFLELTNHFRAFECKDHDSRCKOPSQTLSP
SDFKLKNGRTSGYDARLRPNFKGPPVNTCNIFINSFSEVETMDKRVNLFKQW
NDSRLAYSEYDPDSDLDPSMDSIWKPDLPFANEKANFHDVTDNKLRLSKNGV
LVSIRLTLLSCPMDLKNPMDVOTCTMOLSEFGYTMNDLIEFMLSGCVQVAGEITL
POFLKEKEELGCTKHYNTGKTCIEVFHLEROMGYKTIOMYIPSLIVILSWVF
WINMDAPARVAHIGITTYVTCTTOSGSGSASIPKSYVAIDIMAVCLLPFALLE
YAAVNFASRKHKEFLRLRRQKRNKEEDVTRESNFSYGMGHCLQVKKDSTAVAKT
PAMPPEPRKDDAIKKKFDRAKRIDTISRAPFLAFLENIIFYWITYKIRHEDVH
KK"
BASE COUNT      526 a      416 c      417 g      498 t
ORIGIN
Alignment Scores:
Pred. No.:      4,94e-161      Length:      1857
Score:          1682.50      Matches:      324
Percent Similarity: 86.10%      Conservative: 29
Best Local Similarity: 79.02%      Mismatches: 36
Query Match:      74.71%      Indels:      21
DB:              9      Gaps:      5
US-10-075-846-4 (1-431) x HSGLYRA1 (1-1857)
OY 24 ATGVALAlaLeuAlaLysGlu---GluValLysSerGlyThrLysGlySerGlnProMet 42
DB 454 AGGACGGCTTTCTTCGCAAGAACCATGACCTCCAGTGGAAAACACCTTCACGACCCCTA 513
OY 43 SerProSerAspPheLeuAspLysLeuMetGlyArgThrSerGlyTyrAspAlaArgile 62
DB 514 TCCCTTCAGATTTCTTGACAGATTAAAGGGAAGACATCAGATATGATGCAAGAATC 573
OY 63 ArgProAsnPheLysGlyProProValAsnValThrCysAsnIlePheIleAsnSerPhe 82
DB 574 AGCCCAAAATTTTAAAGTCTCCAGTAAACGTTACTGCAATTTTATCAACACGTTT 633
OY 83 SerSerValThrLysThrMetLaspTyrArgValAsnValPheLeuArgGlnGlnTrp 102
DB 634 GGATCAGTCACAGAAACGACCATGACATCCAGTGAATATTTTCTGACACACAGTGG 693
OY 103 AsnAspProArgLeuSerTyrArgGluTyrProAspAspSerLeuAspLeuAspProSer 122
DB 694 AATGATTACAGCGCTGCGTACGTAGTACCCAGATGATGATCCCTGCGACTGGACCATCC 753
OY 123 MetLeuAspSerIleTrrLysProAspLeuPhePheAlaAsnGluLysGlyAlaAsnPhe 142
DB 754 ATCTGAGACTCCATTGGAAACAGATTTGCTTGGCCAATGAGAAAGGCTGCCAATTC 813
OY 143 HisGluValThrThrAspAsnLysLeuLeuArgIlePhePheAsnGlnAsnValLeuTyr 162
DB 814 CAGAGATGTCACCACTGACACAAATGCTACGGAATTTCCAAAATGGAAGAGTGGCTTAC 873
OY 163 SerIleArgLeuThrLeuIleLeuSerCysLeuMetAspLeuLysAsnPheProMetAsp 182
DB 874 AGTATCAGACTACCTTACCTTATCCCTGCCATGAGACTTGAAGAATCTTCGATGAGAT 933
OY 183 IleGlnThrCysThrMetGlnLeuGluSerSerIleLeuLysSerProLeuProSer 202
DB 934 GTCCAGACCTGATCAATGACGCTGAG----- 960
OY 203 LeuSerLeuSerValGlyTyrThrMetLysAspLeuValPheGlnTrpLeuGluAspAla 222
DB 961 -----AGTTTGGGTACAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1011
OY 223 ProAlaValGlnValAlaGlnGlyLeuThrLeuProGlnPheIleLeuArgAspGluLys 242
DB 1012 CCA---GTCCAGATTTGCTGAAGATTGACCTGCCAGTTTATTTTGAAGAAGAAGAG 1068
OY 243 AspLeuGlyCysCysThrLysThrLysThrAsnThrGlyLysPheThrCysIleGluValLys 262
DB -----
```

```
DB 1069 GAACCTGGCTACTGACAAAGCACTACAAACACTGGAAAGTTTACCTGCATTGAGCTCAAG 1128
OY 263 PheHISLeuGluArgGlnMetGlyTyrTyrLeuIleGlnMetTyrIleProSerLeuLeu 282
DB 1129 TTTCACCTGGAGCGCAATATGGATATTTATTTGATCCAGATATACATCCCAAGCTGCTT 1188
OY 283 ILevalIleLeuSerTyrValSerPheTrrPileAsnMetAspAlaAlaProIleArgVal 302
DB 1189 AATGAAATTTTGTCTCGGTGTTCCCTTTGGATAAATATGATGACAGCCCTCCAGGGTTC 1248
OY 303 GlyLeuGlyIleThrThrValLeuThrMetThrThrGlnSerSerGlySerArgAlaSer 322
DB 1249 GCATCGGGCATCCACACAGCTTAAAGATGACACACAGAGTTACAGCTCCAGCGGCATCT 1308
OY 322 LeuProLysValSerTyrValLysAlaIleAspIleTrrPmetAlaValCysLeuLeuPhe 342
DB 1309 CTGCCAAAGGTTCTCTATGTAAAGCGATTGACATCTGATGCGCTGTGCTTCTGTTT 1368
OY 343 ValIleAlaAlaLeuLeuGlnLysValIleAlaIleAsnPheValSerArgGlnHisLysGlu 362
DB 1369 GTGTTTGCTGGCTTACTGGAATACGAGCGGTGAACCTTGCTCCAGGACACAAAGAG 1428
OY 363 PheIleArgLeuArgArgGlnArgGlnArgGlnArgGlnArgGlnArgGlnArgGlnArg 382
DB 1429 TTCCTGCGGCTCCGAGAGAGACAGAGAGAGCAAGATTAAGAGAGAGACCTTACTGTGAA 1488
OY 383 SerArgPheTyrPheArgGlyTyrGlyLeuGlnLysCysLeuGlnAlaIleArgAspGlyGly 402
DB 1489 AGTCGTTTAAATTTATGACGCTTATGAGATGGTACTGCTCTCAAGTGAAGAATGAGACA 1548
OY 403 PrometGlnGlySerGlyIleTyrSer---ProGlnProProAlaProLeuLeuArgGlu 421
DB 1549 GGTGTCAGAGCCACACACCTGCCAACCCACTCCCAACAGCCCA-----AAAGAT 1596
OY 422 GlyGluThrThrArgLysLeuTyrValAsp 431
DB 1597 GGAGATGCTATCAAGAAAGATTGTGGAC 1626
RESULT 6
LOCUS      RNO310837      1359 bp      mRNA      linear      ROD 03-APR-2001
DEFINITION Rattus norvegicus mRNA for glycine receptor alpha 2 precursor.
ACCESSION  AJ310837
VERSION     AJ310837.1 GI:13548660
KEYWORDS    glycine receptor alpha 2 precursor.
SOURCE      Norway rat.
ORGANISM    Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE   1 (bases 1 to 1359)
AUTHORS    Beato M., Groot-Kormelink P.J., Colquhoun D. and Slyvlotz L.G.
TITLE      Concentration dependence of single channel currents through rat
           recombinant alpha 1 glycine receptors
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 1359)
AUTHORS    Groot-Kormelink P.J.
TITLE      Direct Substitution
JOURNAL     Submitted (30-MAR-2001) Groot-Kormelink P.J., Department of
           Pharmacology, The School of Pharmacy, 29/39, Brunswick Square,
           London, WC1N 1AX, UNITED KINGDOM
FEATURES
   source
   1..1359
   /organism="Rattus norvegicus"
   /strain="Sprague-Dawley males"
   /db_xref="taxon:10116"
   /tissue_type="whole brain"
   /clone_lib="rat brain marathon-ready cDNA"
   /dev_stage="8-12 weeks"
   1..1359
   /function="ion channel"
   /codon_start=1
   /evidence=experimental
```

**THIS PAGE BLANK (USPTO)**

Db	1264	CTCTTTGGTTCGCGCTGCCTTCTGAGATGATGTCATAAATTTTGTTCGTCAGCAT	1323
Oy	1081	AAAGAATTCATACACACTTTCGAAGAAGCAGAGGGCCCAACGCTTGGAGGAAGATATCATC	1140
Db	1324	AAAGAATTCATACACACTTCGAAGAAGCGAGGGCCCAACGCTTGGAGGAAGATATCATC	1383
Oy	1141	CAAGAAGTGGTTCATATTCGCGGGCATGGCTTGGGCCATCGCTGCGACGCAAGGAT	1200
Oy	1201	GGAGGTCCAAATGGAAGGTTCTGGCATTTATAGTCCCAACCTCCAGGCCCTCTCTAAG	1260
Db	1444	GGAGGTCCAAATGGAAGGTTCTGGCATTTATAGTCCCAACCTCCAGGCCCTCTCTAAG	1503
Oy	1261	GAAGGAGAAACCAACGCGGAAACTCTAGCTGAGCTGAGGCAAGAGAAATTGACACATCTCC	1320
Db	1504	GAAGGAGAAACCAACGCGGAAACTCTAGCTGAGCTGAGGCAAGAGAAATTGACACATCTCC	1563
Oy	1321	CGGGCTGTCTCCCTTCACATTTTCTCTCATATCTTCAATATCTTCTACTGGGTGTCTATAA	1380
Db	1564	CGGGCTGTCTCCCTTCACATTTTCTCTCATATCTTCTACTGGGTGTCTATAA	1623
Oy	1381	GTGCTATGTCAGAAATATCCACACAGCTCTGTAATAGGTTGGAGCTTATAGTCTCT	1440
Db	1624	GTGCTATGTCAGAAATATCCACACAGCTCTGTAATAGGTTGGAGCTTATAGTCTCT	1683
Oy	1441	GCTGCTGGCCCTCCCTTCCTCTCTGTTGGTGGCTTTCCTCCTAGTTAGCTTCATTAGG	1500
Db	1684	GCTGCTGGCCCTCCCTTCCTCTCTGTTGGTGGCTTTCCTCCTAGTTAGCTTCATTAGG	1743
Oy	1501	GTTTGGAGATGTCCTTCTCTGATCTCCACTCAGAACTCAACTACCAATCCCAAGCTAT	1566
Db	1744	GTTTGGAGATGTCCTTCTCTGATCTCCACTCAGAACTCAACTACCAATCCCAAGCTAT	1803
Oy	1561	GTGGGCTATATATGTCATGTCGTCGCAATGGATGGCTTATATAAGATGGCTTATACCC	1620
Db	1804	GTGGGCTATATATGTCATGTCGTCGCAATGGATGGCTTATATAAGATGGCTTATACCC	1862
Oy	1621	TAAAAAATAAAAAA 1632	
Db	1863	TAAAAAATAAAAAA 1874	
RESULT 2			
AAC61678			
ID	AAC61678 standard; DNA; 1600 BP.		
XX	AAC61678;		
AC			
XX			
DT	19-FEB-2001 (first entry)		
XX			
DE	cDNA sequence encoding a human ataxia protein.		
XX			
KW	Human; ataxia; gene therapy; ss.		
XX			
OS	Homo sapiens.		
XX			
FH	Key		
FT	CDS		
FT	Location/Qualifiers		
FT	1..1254		
FT	/*tag="a		
FT	/product="ataxia protein"		
XX	WO200058461-A1.		
PN			
PD	05-OCT-2000.		
XX			
PF	23-MAR-2000; 2000MO-EP02600.		
XX			
PR	26-MAR-1999; 99EP-0106343.		
XX			
PA	(RAPP/) RAPPOLD-HOERBRAND G.		
XX			
XX			
PI	Rappold-Hoerbrand G;		

Query Match	Best Local Similarity	Mismatches	Conservative	Score	Pred. No. 0	DB 21	Length	1600
Query 1	91.7%	1585	0	1503.8	0	21	1600	
DB 1	96.6%	12	44	1503.8	0	21	1600	
DB 61	91.7%	1585	0	1503.8	0	21	1600	
DB 121	91.7%	1585	0	1503.8	0	21	1600	
DB 121	91.7%	1585	0	1503.8	0	21	1600	
DB 121	91.7%	1585	0	1503.8	0	21	1600	
DB 181	91.7%	1585	0	1503.8	0	21	1600	
DB 181	91.7%	1585	0	1503.8	0	21	1600	
DB 241	91.7%	1585	0	1503.8	0	21	1600	
DB 241	91.7%	1585	0	1503.8	0	21	1600	
DB 301	91.7%	1585	0	1503.8	0	21	1600	
DB 301	91.7%	1585	0	1503.8	0	21	1600	
DB 361	91.7%	1585	0	1503.8	0	21	1600	
DB 361	91.7%	1585	0	1503.8	0	21	1600	
DB 421	91.7%	1585	0	1503.8	0	21	1600	
DB 421	91.7%	1585	0	1503.8	0	21	1600	
DB 481	91.7%	1585	0	1503.8	0	21	1600	
DB 481	91.7%	1585	0	1503.8	0	21	1600	
DB 541	91.7%	1585	0	1503.8	0	21	1600	
DB 541	91.7%	1585	0	1503.8	0	21	1600	
DB 601	91.7%	1585	0	1503.8	0	21	1600	
DB 601	91.7%	1585	0	1503.8	0	21	1600	
DB 661	91.7%	1585	0	1503.8	0	21	1600	
DB 661	91.7%	1585	0	1503.8	0	21	1600	
DB 721	91.7%	1585	0	1503.8	0	21	1600	
DB 721	91.7%	1585	0	1503.8	0	21	1600	

**THIS PAGE BLANK (USPTO)**



**THIS PAGE BLANK (USP)**



CCC

Pf	compounds useful for treating disorders relating to mutations in ataxia gene -
xx	
PS	Claim 2; Page 18-20; 47Pp; English.
xx	
CC	The present sequence encodes a human ataxia protein. The ataxia
CC	protein and polynucleotides are useful for diagnosing and treating
CC	disorders related to ataxia. Ataxia gene sequences are useful in
CC	gene therapy, and as diagnostic tools or reagents for identifying and
CC	characterizing genetic defect involved in the disorders and diseases
CC	related to ataxia.
xx	
SO	Sequence 1600 BP; 378 A; 440 C; 379 G; 403 T; 0 other;
	Alignment Scores:
	Pred. No.:           4,066-240           Length:           1600
	Score:             2157.00               Matches:          414
	Percent Similarity: 96.29%              Conservative:    1
	Best Local Similarity: 96.06%            Mismatches:      2
	Query Match:       95.78%               Indels:           14
	DB:                   21                   Gaps:            1
	US-10-075-846-4 (1-431) x AAC61678 (1-1600)
Oy	
Db	1 MetThrTrHLeuValProAlaThrLeuSerPheLeuLeuLeuTrpThrLeuProGlyIn 20
	1 AAGCAACACTCTTGTCCTGCACACCCTCTCTTCTCTGTGGACCTGCCAGGGCAG 60
Oy	21 ValLeuLeuArgValAlaLeuAlaLysGluValLysSerGlyThrLysGlySerGln 40
Db	61 GCCTCCTCAGGGGTGGCCTTGGCAAAGAGAAGTAATCTGGAACCAAGGGGTCCAG 120
Oy	41 PrometSerProSerAspPheLeuAspLysLeuMeGlyIArgThrSerGlyTrysPaLa 60
Db	121 CCCATGTCCCCCTCGATTCTCTAGACAACATTATGGGGCAGAACATCGATATGATGCC 180
Oy	61 ArgIleArgProAsnPhenylsGlyProproValAsnValThnCysAsnIlePheIleasn 80
Db	181 AGCATTCGGCCCAATTTTAAAGGCCACCCTGGAGAGTAGCTACGCAACATCTTCATCAC 240
Oy	81 SerPheSerSerValThrLystrHrmElaSPtyrArgValAsnValPheLeuArgGln 100
Db	241 AGTTTCAGCTCCATCACCACGAACCAATGAGACTACGGGTGAATGTCTTCTGGCGCA 300
Oy	101 GlntPrasnaSPProArgLeuSerTyrrArgGluTyrrPrAspaSPserLeuAspLeuasp 120
Db	301 CAGTGAAGAGCCACACGCGCTGCTCCAGCAAGATAATCCGATGACTCTCGACCTCGAT 360
Oy	121 ProSerMetLeuAspSerIleertrylspProAspLeuPhehalaSngIulysgIyala 140
Db	361 CCCTCCATCTGGACCTATCTCGAAGCCAGACCTCTTCTTGCTATATGAAAGGGGCC 420
Oy	141 AsnPheHisGIvAlThrThraspaSnLysLeuLeuArgIllepheLysAsngIyaSnval 160
Db	421 AACTTCATGAGGTGACCAAGGACCAAGATTACTCGCATCTTCCAAGATGGGAATGTG 480
Oy	161 LeutySerIleArgLeuThrLeuileLeuSerCysLeuMetAspleuLysaSnPhePro 180
Db	481 CTGTACAGCATCAGCGTGAACCTCATTTGTCTGCTGATGGACCTCAAGAACTTCCC 540
Oy	181 MetAspIleGlnThrCysThrmetGlnLeuGluSerSerSerIleucysSerProLeu 200
Db	541 ATGGACATTCAGACCCGACAGATGCGAGCTTGAC----- 573
Oy	201 ProSerLeuSerLeuSerValGlyTyrrHmetyAspLeuValPheglutRpleGlu 220
Db	574 -----AGCTTGGCTACACCATTAAGACCTCGTGTGTGAGTGGCTGAA 618
Oy	221 AspAlaProAlaValGlnValalagIugIyleuThLeuProGlnPheIleLeuArgasp 240
Db	619 GATGCTCGTCGTGCCAAGTGGCTGAGGGGCTACCTGCCCAAGTTATCTTGGCGGAT 678
Oy	241 GlulysAspLeuGlucysCysthrLyshIstYrAsnThrclLysSphehrCysIleglu 260

**THIS PAGE BLANK (USPTO)**



**THIS PAGE BLANK (USPTO)**